

Media studies for scientists

Science, with its inherent uncertainties, can be hard to put across to the public. But blaming 'sloppy' journalism is too easy. If researchers are to make their points effectively, they should learn more about how the media work.

No organization likes to feel the pressure of negative media coverage. Last year, when the Fred Hutchinson Cancer Research Center in Seattle was accused of mismanaging clinical trials, this bastion of the medical establishment found itself under an unwanted spotlight. Now the series of articles that made the allegations is a contender for the Pulitzer Prize, and those involved are reliving their media nightmare (see page 463).

Leading researchers and scientific bodies have rallied round the centre. This was bad journalism, say the critics. It seems like a familiar story: sloppy journalists misrepresent science.

How often does this happen? The cancer-centre story is a complicated and emotive one. But there is no doubt that science is, on occasion, badly covered in the media. Last week, for instance, one British newspaper ran a front-page story headlined "Scientists find Prozac 'link' to brain tumours". The lead author of the research told *Nature* that the story was "scandalous" — there is no epidemiological link. His study merely showed that the antidepressant could, in the test tube, prevent certain tumour cells from being triggered to commit suicide by the neurotransmitter serotonin.

But such examples are the exception. In the United States, for instance, media coverage of climate change has been thorough, explaining the consensus view while noting dissenting voices.

In Britain, newspapers, including the one that carried the Prozac story, have produced some excellent science journalism. Much of the coverage of a recent public scare over the safety of the combined measles, mumps and rubella vaccine, for instance, stressed that most experts regard the vaccine to be safe, and pointing out the limitations of the small studies that have linked it to the development of autism. Even the Prozac story qualified the nature of the 'link' further down the piece. Take away the inflammatory headline, and the prominent positioning in the paper, and the story may have caused little fuss.

Nevertheless, many scientists are quick to attack the media when they believe they have been misrepresented. And at conferences addressing the public understanding of science, journalists are often portrayed as the root of the problem.

This knee-jerk reaction itself misrepresents the way in which stories enter the media — journalists are not the only players. University press offices, in particular, must take some responsibility for hyped findings. Pushed by university leaders to maximize total coverage, press officers fill their releases with claims of significant breakthroughs. Scientists then complain when their work is hyped beyond its true worth.

Researchers could also examine how other professions manage the media. Politicians are misrepresented more frequently and significantly than scientists. But they know that attacking journalists is a short-sighted strategy. Instead, they have become experts in rebutting inaccurate stories and imparting their own message.

How far should scientists go down this road? Scientists, as with almost every other profession, enjoy more public trust than politicians. Gaining a reputation for spin could damage this. Researchers should, however, learn what the media wants. Politicians understand the kind of stories that journalists are looking for. If more scientists did too, they would be better equipped to get their message across.

Some progress has already been made in this direction. Many grant-awarding bodies now promote media training for scientists. Britain's Engineering and Physical Sciences Research Council, for example, recently announced plans to include £500 (US\$720) for media training in each grant it awards.

Better-funded initiatives of this sort are needed, and inaccurate reporting should not go uncorrected. But attacks on journalists often sound bitter, and have little long-term benefit. Rather than shooting the messengers, scientists should take them to one side and give them the real story. ■

Prevention remains the best medicine

For the foreseeable future, smoking will kill more people than cancer researchers can hope to cure.

After decades of research into the biology of cancer, it is good to finally see some signs of a therapeutic pay-off (see page 470). But the statistics of cancer mortality still make depressing reading. All the more so given that, in major Western countries, tobacco plays a part in one-third of cancer deaths — mainly, but not exclusively, through lung cancer.

This vast toll underlines the continued importance of prevention — and especially of campaigns to reduce smoking — in the war on cancer. Stopping the relentless promotional activities of tobacco companies would be a major step forward.

But our imperfect world is home to conflicting political and economic forces. In Europe, for example, the European Commission gained some responsibility for public health with the 1993 Maastricht Treaty, and has an active anti-smoking policy. Yet the European Union spends around 1 billion euros (US\$900 million) each year subsidizing tobacco

farmers, three-quarters of whom are in Italy and Greece; Spanish and French farmers receive most of the remaining subsidies. In total, the handouts are broadly comparable to the amount spent annually on cancer research in Europe's public sector. The European Commission wants to phase out the subsidies, which have their historical roots in the European Union's outmoded Common Agricultural Policy, but is facing resistance from some of the member states that benefit.

Researchers are not above these messy conflicts. A proposed European Union directive to restrict tobacco advertising is now being hotly debated, while scientists continue to accept the patronage of Philip Morris, the world's largest tobacco company, which awards a prestigious and generous research prize in Germany every year (see *Nature* **410**, 725; 2001).

Given the centrality of the war on cancer to public health, we should all be asked: which side are you on? ■





Library launch
Online tumour bank
offers easy access to
tissue samples
p464



Bedside manner
Charity tackles
developing world's
drug needs
p465



Powered down
Anger as plug is
pulled on Vivitron
accelerator
p466



Soft landing
India joins the
transgenic cotton
club
p468

Papers square up over potential Pulitzer for cancer-centre critics

Erika Check, Washington

Two American newspapers are fighting a messy public battle over a series of articles that criticize one of the most respected cancer centres in the world.

The Seattle Times ran the series, called "Uninformed consent", last March. The series alleged that the Fred Hutchinson Cancer Research Center in Seattle, Washington, failed to tell patients about the true risks of some experimental cancer trials that it ran in the early 1980s. The series concluded that researchers with financial interests in the trials persisted with their research even after they knew it was killing patients. The centre maintains that there was no financial conflict of interest and said in a statement that "the central themes of *The Seattle Times'* articles are false and unsupportable".

After bagging a slew of awards, "Uninformed consent" is now rumoured to be a finalist for a Pulitzer Prize, the premier award in American journalism. But with the Pulitzer



Under attack: the Fred Hutchinson Cancer Research Center was heavily criticized by *The Seattle Times*.

board set to announce the prizewinners on 8 April, *The Wall Street Journal* has launched what observers say is an unprecedented attack on *The Seattle Times* series.

The episode recalls this year's Oscar fight, when Hollywood studios blamed reporters for running a smear campaign against the film *A Beautiful Mind*. But although the Oscars are known for inspiring massive lobbying crusades, journalists say that they have never seen anything like this unruly Pulitzer squabble.

The battle began on 19 March, when Laura Landro wrote an opinion column in *The Wall Street Journal* slamming "Uninformed consent". Landro, an assistant managing editor at *The Wall Street Journal*, called the series "fundamentally false".

"Rather than racking up prizes, it should be used as a textbook case on how the media can convey biased and misleading information about biomedical research," Landro wrote.

Landro noted, for example, that the Hutchinson centre was one of the first cancer centres in the world to acknowledge problems with a trial involving the removal of T cells from bone marrow donated in transplants, when problems caused by the removal became known.

Michael Fancher, executive editor of *The Seattle Times*, leapt to his paper's defence on 22 March. He accused Landro of failing to disclose her own conflict of interest. Fancher said that Landro had donated money to the Hutchinson centre, widely nicknamed 'the

Foresters cautious over transgenics

Virginia Gewin, Washington

American farmers may have jumped at the chance to use genetically modified crops, but foresters are set to take a more cautious approach.

That is the main message of a report on genetically engineered trees just published by the Washington-based Pew Initiative on Food and Biotechnology.

Forest biotechnology has the potential to create pest-resistant trees, to design trees for more easily processed wood products, and possibly to restore endangered species.

But environmental, regulatory and societal issues, which have been problematic for agriculture, could spell more trouble for forestry, says the report, which was written on the basis of a conference run by the Pew Initiative in December. Because trees can spread their pollen over long distances, and plantations are often close to natural forests, gene flow into wild populations would be

inevitable, the report says.

The meeting was co-sponsored by the Society of American Foresters (SAF) and the Ecological Society of America, and sought to bring ecologists and foresters together to identify key issues in the debate. Such issues include risk assessment, public participation in decision-making, and possible changes to regulatory oversight of forest biotechnology.

"You just can't afford to make an investment in forestry that's too risky, because it takes so long for the trees to grow," says Terry Clark, science manager with SAF.

But Toby Bradshaw, a forestry expert at the University of Washington in Seattle, says that transgenics will be important in forest research even if it is not used in commercial forestry.

The Pew Initiative hopes to take recommendations for improving US policies and programmes related to forest biotechnology to Congress by this autumn. ■

Hutch', after she was successfully treated there in 1992.

"Unfortunately, Ms. Landro is unable to separate her own experience as a patient from her duties as a journalist," Fancher wrote in *The Seattle Times*.

From there, the tussle escalated, even though editors at both papers say that *The Wall Street Journal* is not competing with "Uninformed consent" for a Pulitzer in the investigative reporting category.

On 25 March, *The Wall Street Journal* published a letter from Fancher in which he accused the Hutch of failing to enact reforms of its clinical-trials oversight process that had been recommended by a review committee. He also said the Food and Drug Administration had found problems with the centre's informed-consent process and shut down three clinical trials there in June 2001.

The Wall Street Journal then devoted an unsigned editorial to an attack on "Uninformed consent" on 27 March. The editorial accused *The Seattle Times* of sexism and ambulance-chasing. It implied that the paper had not published a previous letter from Landro because the letter would have had to be included in the Pulitzer submission. It also ran a series of letters from scientists and doctors criticizing *The Seattle Times* series.

On *The Wall Street Journal's* editorial page, editors claimed that far more is at stake than just a prize for journalists. "We have long felt that medical research needs to be liberated, not inhibited," they wrote.

Some science journalists say that they are concerned that *The Seattle Times* series was not as thorough or as scientifically rigorous as it should have been.

B. D. Colen, winner of a Pulitzer Prize and faculty member of the graduate programme in science writing at the Massachusetts Institute of Technology, says that the first instalment of the series did not give complete information about the different stages of clinical research, or where the Hutch's experiments fitted into these stages.

But despite the criticism levelled at the series, some observers are defending it on the grounds that the award of the Pulitzer would encourage more investigative reporting into science and medicine.

Another Pulitzer winner, Deborah Blum, president-elect of the National Association of Science Writers, says that she has not read the series. But she welcomes it, saying that there is too little investigative journalism into science and medicine.

"It does not bother me at all that a local newspaper made a cancer centre angry in its investigation," Blum says. "Especially with clinical trials, it's such a good thing to get inside them and try to make people understand what they are about." ■

Online tumour bank aims to offer ready route to tissues



Researchers fear that current access to tumour samples is not sufficient to unpick cancer's secrets.

David Adam, London

It is a pathological paradox: how to meet cancer researchers' increasing demand for human tissue while public concern about its use reduces the amount available.

Last week the European Organisation for Research and Treatment of Cancer (EORTC) unveiled its solution — an online library of tens of thousands of tumour tissue samples held at research centres and tissue banks across Europe.

Under plans discussed at last month's EORTC conference in Brussels, participating researchers would browse through samples online and take enough tissue and associated clinical data to carry out molecular profiling and clinical trials. The frozen samples would be stored at the respective centres, and a representative section then scanned and entered into the virtual library.

"By making more tissue more easily available to scientists we can speed up translational research tremendously," says Wolter Oosterhuis, a pathologist at University Hospital Rotterdam who is leading the project.

An increase in both regulation and public scrutiny of the use of research tissue is making access to some materials increasingly difficult. Nick Lemoine, a cancer researcher at London's Hammersmith Hospital, says that supplies of tissue "almost dried up" in some parts of Britain following a series of high-profile scandals over retention of human tissue without consent. Others say that a shortage of pathologists and a tendency to take smaller amounts of material during each biopsy have made some tumour tissue harder to obtain.

The plans are at an early stage, but Oosterhuis is hopeful that the system can be up and running within four years. It builds on a similar EORTC initiative to digitize thousands of slides of tissue samples to create a virtual tumour-tissue bank. Such slides allow physical comparison between samples, but fresh tissue is more useful as it allows investigation of DNA, RNA and protein expression, which researchers believe will shed light on genetic contributions to cancer (see page 470).

Eight cancer centres in France, the Netherlands, Spain, Austria and Britain have signed up to the scheme, although some centres have refused to take part. "Scientists are not easily separated from their materials," Oosterhuis says, "but we argue that by making it available to others you get access to many more samples than you can ever collect yourself."

The move comes as some researchers express concern that the existing network of tissue banks cannot meet the growing demand for materials. "There has been a huge expansion of molecular analysis on the basis of the genome and we need to provide the materials for this tissue analysis to be done," says Mike Stratton, who heads the Cancer Genome Project at Britain's Sanger Centre near Cambridge.

Some countries, such as the United States and Spain, already link the contents of local banks into a national network. Others are setting up centralized national facilities for tumour tissue. In Britain, a generic tumour-tissue bank that will ultimately contain some 3,000 samples from 23 different cancers is expected to open at the University of Glasgow by the end of this year. ■

Charity launches not-for-profit drug industry

Declan Butler, Paris

The Paris-based charity Médecins Sans Frontières (MSF) has launched plans for an ambitious international effort to research, develop and produce drugs for neglected diseases that afflict people in poor countries.

The Drugs for Neglected Diseases Initiative (DNDI) will do “nothing short of creating a global, not-for-profit pharmaceutical industry”, says Philippe Kourilsky, director-general of the Institut Pasteur in Paris.

The initiative is being undertaken because the charity — known in the United States as Doctors without Borders — is finding its efforts to deliver medical care increasingly hampered by expensive, ineffective or unavailable drugs, MSF officials say.

The DNDI was formally launched in New York last month, at a meeting attended by 400 public-health experts, among them Harold Varmus, the former director of the US National Institutes of Health (NIH), and Jean-Pierre Garnier, chief executive of the multinational drug company GlaxoSmith-Kline. The project is being initially backed by the Pasteur, Brazil's Oswaldo Cruz Institute, the Indian Council of Medical Research, the Science University of Malaysia and the World Health Organization.



Overcoming neglect: workers in developing countries complain of a lack of available treatments.

MSF, which won the 1999 Nobel peace prize for its worldwide public-health efforts, has put up an initial \$1 million for five pilot projects to develop and produce drugs for the parasitic diseases visceral leishmaniasis and sleeping sickness. And the DNDI will conduct a one-year study to determine the shape

of the larger initiative. “This will hammer out realistically what we can and cannot do,” says Kourilsky. More funding will then be sought from foundations and governments — \$50 million annually may be sufficient, says Dyann Wirth, a tropics expert at Harvard School of Public Health.

Established drug companies don't develop drugs for diseases with poor commercial prospects, says Bernard Pecoul, a Geneva-based MSF official — hence the need for the DNDI. He says the initiative's goal is to create a virtual industry driven by “public need” rather than share price.

Wirth says that the idea is not to start up an integrated operation from scratch, but instead to coordinate an international network of researchers, institutions and drug manufacturers. Existing drug companies in developing countries such as India and Brazil will probably have a major role, Wirth adds.

Just as MSF has mobilized doctors worldwide, Pecoul is keen to “mobilize” the research community worldwide. “If we hear of interesting academic work, the network will put them in touch with others to take projects forward,” he says.

Big drug companies, which previously clashed with MSF over its efforts to supply low-cost AIDS drugs to Africa, are welcoming the initiative, at least in public. Garnier pledged his support at the New York meeting. Industry has been “very open”, says Pecoul, although contributions such as access to molecular-data libraries and the lifting of patent rights will be on a case-by-case basis.

Gerald Keusch, director of NIH's Fogarty International Center, describes the project as “hugely ambitious”. But he suggests that the time might be ripe for such an undertaking, and says the Fogarty centre is in discussions with MSF about becoming involved.

Putin reads science the riot act

Bryon MacWilliams, Moscow

Russian Federation president Vladimir Putin has issued an unusually blunt assessment of the state of Russian science — and pledged that government funding will be swiftly redirected to the most commercially viable disciplines.

“Today, governmental support of science is completely ineffective,” Putin told a rare meeting of the most senior officials in the government and its scientific advisers at the Kremlin on 20 March.

“Everyone is claiming to be on the path of innovation — but almost nothing has been done in real terms,” Putin said, adding that Russian science was poorly adapted to the market and over-dependent on government funds. “The business sector cannot find a common language with science,” he said.

Putin pledged, nonetheless, to put more government money into research and development. But much of it will be directed at technology development, rather than basic research. “This is the first step towards a sensible, self-regulated departure from the senseless scattering of resources” of the past, he said.

Government officials at the meeting said



Russia's president: the business sector has no common language with science.

that, under a reform plan to be fully implemented by 2010, the existing network of scientific institutes will be pruned, and resources sent instead to areas regarded as critical to Russia's interests. They claimed that funding would be increased by a factor of five in areas including telecommunications, electronics, aviation, new materials and chemistry.

The president also pledged to raise salaries for young scientists. Although observers are sceptical of his ability to deliver on this and other aspects of the reform plan, scientific leaders welcomed his participation at the conference.

Sparks fly as electrostatic facility fizzles out

Declan Butler, Paris

European nuclear physicists are up in arms over a decision to shut one of the world's largest van de Graaff electrostatic accelerators, the two-million-volt Vivitron in Strasbourg, France.

The researchers say that the move is premature, as the field's next generation of machines is still on the drawing board. They fear that the measure will drive some Europeans who study 'exotic' nuclei out of the field for good.

The study of exotic nuclei — which have an unstable balance of neutrons and protons, and so are prone to rapid radioactive decay — is one of the hottest areas of contemporary nuclear physics. The research is expected to yield clues as to how elements form in stars and supernovae. It could help to refine the standard model, and may even give rise to better medical isotopes.

The Vivitron is one of three facilities in Europe that generate beams of stable ions, alongside machines at Legnaro in Italy and Jyväskylä in Finland. But a more recent technology uses beams of radioactive ions, and there is a consensus in the United States, Europe and Japan that building next-generation machines that use this approach is the top priority for studying exotic nuclei.

The plan to close the Vivitron was announced by the National Institute of Nuclear and Particle Physics, part of France's national research agency, the CNRS. Daniel Guerreau, the institute's deputy scientific director, says that the decision was a simple budgetary one. "Radioactive beams are considered the priority," he says.

But some researchers dispute the idea that existing stable beams are obsolete. They accept that, for most work on exotic nuclei, stable beams will eventually be unable to compete with radioactive-beam technology. But the first radioactive-beam machines, such as the recently opened SPIRAL in France (see *Nature* **416**, 114; 2002), generate beams of limited intensity, and will not fulfil the requirements of some experiments.

Diverse experimental needs mean that even if the next generation of high-intensity machines is built, researchers say that a mixture of machines will still be needed. "Sometimes you have a question where you need a radioactive beam, whereas sometimes you need a stable beam," says Herbert Hübel, a physicist at the University of Bonn.

Moreover, although planned facilities such as the Rare Isotope Accelerator proposed at the Argonne National Laboratory in Illinois, the Eurisol project, and another machine at RIKEN in Japan, would combine many of the best features of existing accelerators, they will not come online for at least eight years.

But Daniel Huss, director of the Institute of



Unstable issue: there is disagreement over whether the Vivitron's stable-beam technology remains useful.

Subatomic Research, which houses the Vivitron, is unmoved by such protests. "I entirely agree with the decision," he says. "The physics we do on Vivitron is no longer a priority."

But the move has come under fire from Juha Äystö, chairman of the Nuclear Physics European Collaboration Committee (NuPECC), which meets under the auspices of the European Science Foundation to produce five-year plans for nuclear physics, similar to those created by the US Nuclear Science Advisory Committee (see below).

Äystö says that researchers cannot simply

move elsewhere, as beam time in Europe is already oversubscribed. "There is still very important physics to do on stable beams," he says. NuPECC's next plan is due to be published at the end of this year, and Äystö says that he is "disappointed" that the CNRS has taken its decision without waiting for it.

Several hundred Vivitron users from outside France are now planning to challenge the closure, which they claim will hit students particularly hard. "I have eight PhD students," one of them complains. "Where are they going to get the data to complete their theses?" ■

US physicists unite behind big ideas

Geoff Brumfiel, Washington

Nuclear physicists in the United States are rallying behind a plan to build two major new facilities in the next 10 years.

The two proposed facilities, for research on rare isotopes and on neutrinos, respectively, are backed in a strategic plan just endorsed by the US government's Nuclear Science Advisory Committee (NSAC). But the committee says that the facilities should not go ahead unless extra money is made available for their construction — and this is cash that even committee members acknowledge will be hard to come by.

The Rare Isotope Accelerator (RIA), the panel's top-priority new facility, would allow nuclear physicists to observe 'exotic' isotopes. The RIA would be more flexible than existing isotope accelerators, its advocates say. Heavy ions would be

accelerated into fixed targets at high energies, where their impact would produce a range of rare, short-lived isotopes that could be separated and studied. Researchers think that a better knowledge of such isotopes will further the understanding of how heavy elements form inside supernovae.

The second facility backed by the panel, the National Underground Science Laboratory (NUSL), would study the almost-undetectable particles called neutrinos. The laboratory would be buried some 7 kilometres below ground, to shield its detectors from cosmic rays and other unwanted signals. Currently, US researchers interested in neutrinos must work at underground detector facilities in Japan, Canada or Europe. Neutrino research is part of the US nuclear physics programme for historical reasons.

University research to touch base with NASA

Tony Reichhardt, Washington

NASA centres are being told to reach out to universities in a bid to sharpen their research edge.

As a first step, the space agency is seeking a Californian university to conduct multidisciplinary research and engineering under contract to its Ames Research Center, near San Francisco.

The arrangement is expected to form part of a wider drive by NASA to contract out more of its research and engineering to universities and private companies. New NASA administrator Sean O'Keefe, and the White House Office of Management and Budget (OMB), where he used to work, have each made a public priority of "competitive sourcing" of government activities.

NASA is planning to set up a University Affiliated Research Center (UARC) near Ames, a type of arrangement pioneered by the Department of Defense in facilities such as the Applied Physics Laboratory at Johns Hopkins University in Baltimore, Maryland, in which the centre performs specific research tasks under contract to the sponsor.

The planned centre would focus on areas such as information technology, biotechnology, astrobiology and nanotechnology. Located nearby in a still-to-be-developed NASA Research Park, the UARC would work in close collaboration with what the project's planners call a "lean, civil-service-based core-research centre" at Ames.

NASA hopes that the mix of university faculty, students and NASA researchers at the centre will draw new talent and stimulate



Drawing on expertise: NASA's university-linked centre will be part of its planned research park.

multidisciplinary research and education.

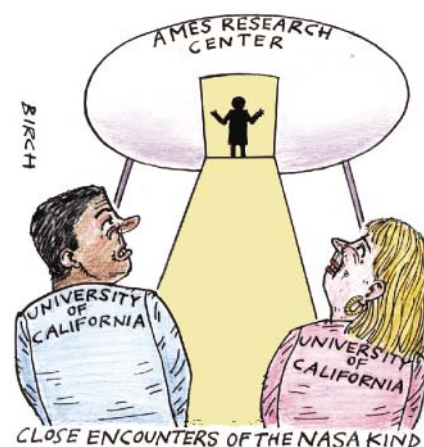
The UARC contract would differ in size and scope from NASA's arrangement with the California Institute of Technology to operate the Jet Propulsion Laboratory (JPL) in Pasadena, California. Whereas the much larger JPL, which has an annual budget of around \$1 billion, is encouraged to be entrepreneurial and to seek work from sources other than NASA, the UARC would be more tightly bound to its sponsor, and outside work would be limited.

Ames asked for expressions of interest from universities in February. If the idea moves forward as planned, a formal solicitation for a UARC operator would go out in August, with a contract to be awarded next February. The budget for the UARC is estimated at between \$10 million and \$20 million for the first year, and \$100 million over a five-year period.

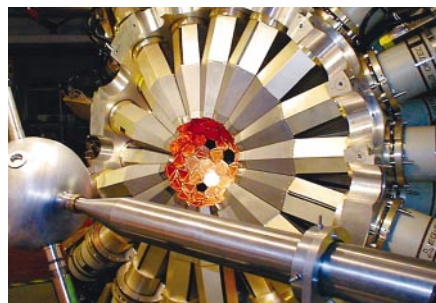
NASA and other federal agencies are under pressure to identify jobs that can be transferred to the private sector. In documents accompanying NASA's 2003 budget

request earlier this year, the Bush administration scolded the agency for not going far enough in this regard. The documents identified the Ames UARC as a "pathfinder effort" in outsourcing, and said future efforts "may include consolidating some NASA facilities with military installations."

NASA tried outsourcing some of its in-house research activities in the mid-1990s, by creating "science institutes" affiliated with universities, but had little success (see *Nature* 380, 7; 1996). Personnel and management issues have proved particularly thorny whenever the subject of privatizing government research jobs is raised. NASA scientists who transferred last year from the Marshall Space Flight Center in Huntsville, Alabama, to a new National Space Science and Technology Center affiliated with Alabama universities have retained their civil-service status and benefits. ■



CLOSE ENCOUNTERS OF THE NASA KIND



Gamma-ray detectors, such as this one at the Berkeley lab, can aid in the study of rare isotopes.

But members of NSAC concede that it will be tough finding money for the two facilities. With its cost estimated at \$700 million, the RIA would put enormous pressure on the physics budgets of the Department of Energy and the National Science Foundation (NSF).

Even so, says James Symons, chair of the committee and a nuclear physicist at Lawrence Berkeley National Laboratory in California, the country should start thinking about the project now. In the end, Symons thinks, any decision to build the RIA will depend on how the proposal stacks up against desired big projects in other sub-disciplines, such as the Next Linear Collider being sought by high-energy physicists.

The NUSL proposal is associated with a site at the Homestake mine in South Dakota (see *Nature* 415, 105; 2002), and has become an issue in a hotly contested campaign for November's election of a South Dakota senator. But the NSF has responded cautiously to suggestions that it should help to build the NUSL, as this would strain its limited budget for large facilities. "There are great political machinations involved," says

NSAC member Alice Mignerey, a nuclear physicist at the University of Maryland. "But we were very careful that it was the science that drove the recommendation."

Top of the priority list in the 10-year plan, however, is full-capacity operation of the two largest existing nuclear-physics facilities — the Relativistic Heavy Ion Collider at Brookhaven National Laboratory in New York state, and the Continuous Electron Beam Accelerator Facility at the Jefferson Laboratory in Virginia. A proposed 6% increase in next year's nuclear-physics budget at the Department of Energy would allow that goal to be met immediately, if it is approved by Congress (see *Nature* 416, 251; 2002). Symons believes that nuclear physicists would then be poised to make some exciting discoveries. "I think this is going to be a great decade for our field," he says. ■

Biotech firms applaud India's acceptance of transgenic cotton

New Delhi India formally opened its doors to genetically modified crops last week. Farmers can now grow transgenic cotton containing a gene from the bacterium *Bacillus thuringiensis* (*Bt*), which produces a toxin that enables the plant to resist attack by its main pest, the bollworm.

Representatives of the biotechnology industry have lobbied intensively for the licensing of the cotton in India, and a counter-campaign by environmental groups has seen the burning of crops at transgenic trial sites. The approval, which was given on 26 March, will allow commercial planting for a three-year trial period. Farmers will be required to grow non-*Bt* cotton around the transgenic variety as a buffer to prevent bollworms from developing resistance to the *Bt* toxin.

The decision was good news for Monsanto, the biotechnology company based in St Louis, Missouri, which owns the rights to the *Bt* gene. But Gene Campaign, an environmental group based in New Delhi, questioned the move. "Indian farms average less than a hectare in size," says a spokesperson. "Can you regulate pollen



Rich pickings: India's move is a boon for biotech companies, but environmentalists are concerned.

dispersal and ensure farmers also retain the right not to grow transgenic crops?"

Police report may herald Porton Down prosecutions

London Charges could be brought against former scientists involved in the British government's chemical and biological weapons research programme, which dates back to the 1950s, according to police reports submitted to the government.

Local police have concluded a three-year investigation into military experiments carried out at Porton Down in Wiltshire. Some servicemen who volunteered for experiments claim that they were told the

research was on the common cold, and many blame the trials for subsequent health problems. The police have sent their results to the Crown Prosecution Service, which will now decide whether to proceed with charges.

Wiltshire police would not confirm whether the report recommends prosecutions, but sources within the force have been reported as saying that the investigation concluded that charges should be brought against three individuals. The police say that those who were directly involved in the most infamous event at Porton Down — the death of airman Ronald Maddison in 1953 — have since died. Maddison died after the nerve agent sarin was allegedly dripped onto a patch taped to his arm.

Lawsuit looms as jobs boycott nears its end

San Francisco The three US national nuclear-weapons labs — Lawrence Livermore in California, and Los Alamos and Sandia in New Mexico — are nearing agreement over ending a jobs boycott by Asian-Americans.

The boycott, which was called in March 2000 by the group Asian Pacific Americans in Higher Education, was in protest at the treatment of Wen Ho Lee, a Chinese-American employee at Los Alamos suspected of spying. Lee was held in solitary

confinement for nine months before all but one minor charge were dropped (see *Nature* 412, 10–11; 2001).

But the Livermore lab may nevertheless be sued by up to 300 current and past employees who allege that it discriminates against Asian-Americans when it comes to salary, hiring and promotion. Nine of them filed a suit against the lab on 22 March, and asked the court to approve it as a class action that would include any Asian-American employee who feels unfairly paid.

Bush offers long-awaited nomination for top NIH job

Washington The waiting is over. President George Bush has officially nominated Elias Zerhouni for the post of director of the National Institutes of Health. If approved, Zerhouni, a radiologist and high-level administrator at Johns Hopkins University in Baltimore, Maryland, will take up a post that has been vacant for over two years.

Materials Update

News about research and applications, 'nanozone', 'material of the month', and content from Nature publications is now available from materials.nature.com

The nomination had been expected for several weeks (see *Nature* 416, 113; 2002). Bush cited Zerhouni's "strong management skills", which he said would be needed to direct the organization's continued expansion. The agency's budget is due to reach US\$27 billion in 2003, double what it was in 1998. Bush also indicated that Zerhouni shared his views in opposing research on therapeutic cloning.

The president's nominee for Surgeon General, Richard Carmona, was also named on 26 March. Carmona is a trauma surgeon at the University of Arizona, Tucson, who serves as a part-time policeman. Bush said Carmona would focus on bioterrorism, healthy living, and alcohol and drug abuse.

Sputnik's launchpad is consigned to history

Moscow Russia is to end most of its satellite launches from the Baikonur Cosmodrome in Kazakhstan, the site from which Sputnik 1, the first ever artificial satellite, was sent into orbit in 1957.

The country's military satellites, and most of its civilian ones, will now be launched from the Plesetsk Cosmodrome in northwest Russia and the standby launch centre, Svobodny, in the east of the country. The move, which will save money by cutting the rent paid to Kazakhstan for the use of Baikonur, should be

complete by 2005.

Plesetsk has already been used for several missions this year, including the launch of the GRACE satellites, a joint project between NASA and the German Aerospace Center to map variations in the Earth's gravitational field (see *Nature* 416, 10–11; 2002).



Russia is phasing out the Baikonur site.

AP

Protesters win revision of German university law

Munich The German government has modified its new university law to placate young scientists who had objected to it (see *Nature* 415, 257–258; 2002).

The law modernizes the rigid employment practices of German universities. But it also restricts the time that scientists can be employed on fixed-term contracts to a total of 12 years. Critics argued that the rule threatened to end the career plans of young scientists who are approaching the limit.

The modified law includes a transition period, which allows researchers who were hired before the bill came into effect to ignore the limit until February 2005.

On the offensive

After decades of disappointment, and the investment of billions of dollars, is the 'war on cancer' about to gain real momentum?

Alison Abbott sends a dispatch from the front line.

Every few years a claim is made for a 'miracle drug' that will cure cancer. But time after time, compounds that have performed wonders in mice have failed miserably when faced with clinical reality.

Even cynics, however, have been taken aback by the performance of a drug called Gleevec, produced by the Swiss company Novartis, and approved by the US Food and Drug Administration (FDA) last year. It is not a cure-all, by any means. But against two particular types of cancer, Gleevec has achieved unprecedented results. For cancer researchers, the drug's remarkable success confirms that they are on the right track: understand which genes go wrong in cancer, design therapeutics to correct these defects, and the disease can be beaten.

"Gleevec is the proof of principle that the

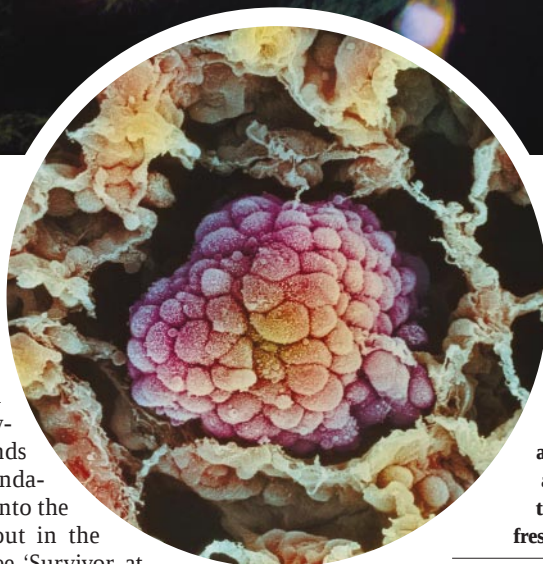
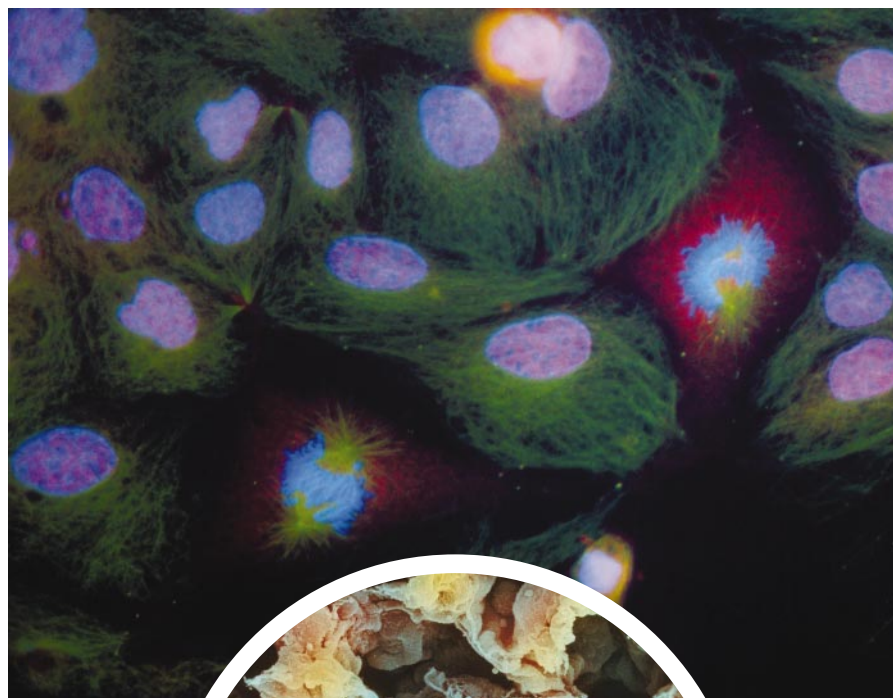
molecular strategy works," says Richard Klausner, until last September director of the National Cancer Institute (NCI) in Bethesda, Maryland, which funds most of the fundamental research into the disease carried out in the United States (see 'Survivor at the helm', opposite).

That proof is long overdue. Conventional cancer chemotherapy, after decades of fine-tuning since it was introduced in the 1950s, has turned around the dismal outlook for

childhood leukaemias — up to 90% of cases are now curable. But against the big cancer killers — including breast, lung, prostate and colon cancers — there has been little progress (see figure, overleaf); sufferers usually experience only a brief period of remission. Even then the price is high, because the drugs are so toxic. Most current chemotherapy agents target dividing cells, for example by blocking the synthesis of new DNA required for cell division — and so hit many healthy organs as well as tumours. In particular, they damage bone marrow, where blood cells are produced.

With more than \$46 billion spent on cancer research by the US federal government alone since President Richard Nixon launched his 'war on cancer' in 1971, a minority of experts has even begun to suggest that cancer has become science's Vietnam. In a cutting essay in the February issue of *Prospect* magazine, for instance, cancer surgeon Michael Baum of University College London claimed that the fight against the disease was bogged down by "slavish adherence to outdated paradigms".

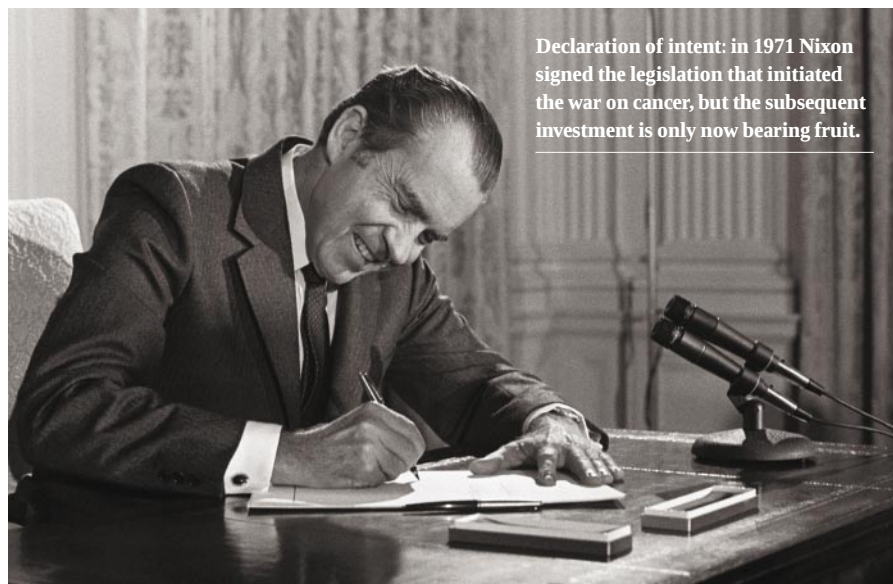
Until Gleevec, promises of kinder and



Growing wrong: despite decades of research, cancer remains a major killer. But improved understanding of the molecular pathways that allow malignant cells, such as these colon (above) and lung (inset) cells, to run riot could offer fresh targets for therapy.

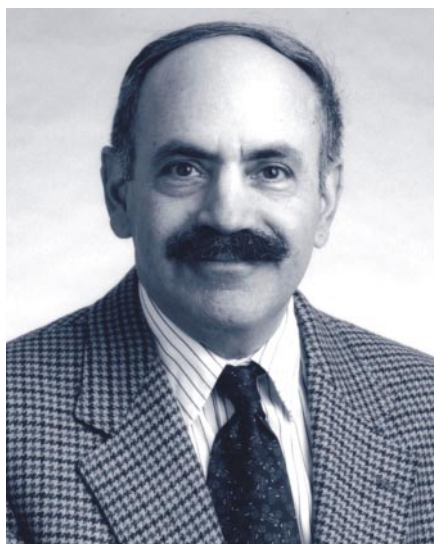
N. KEDERSHA/SPL

MOREDUN ANIMAL HEALTH/SPL



Declaration of intent: in 1971 Nixon signed the legislation that initiated the war on cancer, but the subsequent investment is only now bearing fruit.

BETTMANN/CORBIS



Robert Weinberg believes cancer's complex molecular interactions will give way to simplicity.

more effective therapies had proved empty. But 95% of cases of chronic myelogenous leukaemia (CML) respond to the drug, with the cancer being completely eliminated half of the time^{1,2}. "It was extraordinary for clinicians to see such rapid and dramatic results," says Brian Druker, director of the Leukemia Center at the Oregon Health and Science University Cancer Institute in Portland, who conducted many of the clinical trials. These stunning results, combined with a growing realization among cancer researchers that they have to start delivering the goods, have altered the outlook. The current catchphrase is 'translational research' — aiming to convert molecular insights into effective drugs.

Nixon launched his war on cancer as a successor to President John F. Kennedy's 1960s dream of putting a man on the Moon. But according to some experts, the goal of defeating cancer by 1980 was always unrealistic. "We now understand that cancer is not a simple target that can be approached with high-tech hardware alone," says Klausner. Cancer, it is now realized, is a wily, shifting target — a battery of many different diseases, with a range of underlying causes.

In the vast majority of cases, cancer is an

acquired genetic disease. Cells accumulate a series of mutations that allow them to escape, with ever-greater freedom, the body's normal constraints on their proliferation. "But even mutations promoting unchecked growth are not actually enough to fire a fully fledged cancer," says Douglas Hanahan, a biochemist at the University of California, San Francisco.

In a recent review article³, Robert Weinberg of the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts, together with Hanahan, described six 'hallmarks' of cancer — the acquired characteristics needed to turn a few wayward cells into an aggressive tumour.

Marks of malignancy

First, the cells must mutate so that they can dodge the cellular signals that suppress growth. Then they must acquire their own growth-signalling pathways, independent of external signals. Cells must also evade apoptosis, the system of programmed death under which abnormal cells trigger their own destruction. And they have to develop limitless potential to proliferate: normal cells can divide only about 70 times before their telomeres — the protective caps at the end of chromosomes — become so shortened that the chromosomes are damaged and the cell dies. But cancer cells exploit an enzyme called telomerase to rebuild their telomeres and so escape this constraint.

The other two cancer hallmarks apply only to solid tumours. Growing tumours must create their own networks of blood vessels to deliver the food and oxygen they need. This complex affair, known as angiogenesis, requires a multitude of special growth factors. Finally, the most dangerous tumours are those that have developed mechanisms to allow cells to detach from the main tumour and enter the bloodstream or the lymphatic system. From here they can reach distant tissues, where they grow into secondary tumours, or metastases. Nine out of ten cancer deaths result from metastases.

Very few cancers are caused by a single mutation. "There are so many things that need to go wrong, so it is not surprising that,

in a lifetime, cancer is actually rare," says Weinberg. This is why tumours caught early are usually easier to treat: mutations tend to accumulate as the cancer progresses. The fact that childhood leukaemias are caused by relatively few mutations also helps to explain why they have proved amenable to chemotherapy.

But the multiple changes needed to initiate and support a solid tumour offer a wide potential source of specific targets for scientists trying to develop anticancer drugs. First to be studied were the proteins produced by oncogenes — genes that, when activated,

Survivor at the helm



Taking control:
Andrew von Eschenbach.

After a quarter-century's service in the war on cancer, Andrew von Eschenbach finds himself in a central command post. A surgeon specializing in prostate cancer, the new director of the National Cancer Institute (NCI) in Bethesda, Maryland, declares a "sense of urgency" about the task in hand — he has fought a personal battle against prostate cancer and melanoma. "I know what it's like to wake up in the

middle of the night in a cold sweat and wonder if you're going to make it," von Eschenbach says.

That explains his commitment to the new mantra of 'translational research'. Von Eschenbach pays tribute to his predecessor, Richard Klausner, who ensured that the NCI made major contributions to our understanding of cancer biology. Now, says von Eschenbach, it is time to convert those advances into therapies. "I want to emphasize our applications of that knowledge so that patients' lives are saved, and patients' pain and suffering are relieved," he says.

Like Klausner, von Eschenbach is a strong believer in the molecular approach, and illustrates its potential with an example from his agency's research portfolio. Just a few weeks ago, he notes, NCI scientists revealed that they could diagnose early ovarian cancers by analysing the proteins present in women's blood samples⁴.

But in his new position, charting the future of cancer diagnosis and therapy must sometimes take second place to immediate crises. Von Eschenbach started his job on 22 January; within a month, the NCI had to comment on a public row over the value of breast-cancer screening (see *Nature* **415**, 567; 2002). "While I was still trying to figure out where the restroom was, the mammography controversy landed on my desk," he says.

Von Eschenbach spent most of his career working in the clinic at the University of Texas M. D. Anderson Cancer Center in Houston. And he still intends to spend half a day each week seeing prostate cancer patients. "I want to be able to sit across the bed from a patient and understand the reality of cancer," he says. Erika Check, Washington



Pill power: Gleevec's success has boosted hopes for the molecular approach to cancer therapy.

► promote cell growth and division. Many, such as *ras*, which in 1982 became the first human oncogene to be cloned⁴⁻⁶, do this by stimulating signalling pathways normally activated by growth factors such as epidermal growth factor (EGF). In about 30% of human cancers, *ras* is mutated so that it is permanently switched on, providing a constant growth signal to the cell⁷.

Unfortunately, the discovery of *ras* did not translate into a major clinical advance. Inhibitors targeting the *ras* system were developed by several pharmaceutical companies, but the first generation did not perform well in clinical trials. "We were too excited," admits Mariano Barbacid, director of the Spanish National Cancer Center (see 'The real deal in Madrid', below), who led one of the three groups that independently cloned the gene. "Some people had thought we had opened the door to curing cancer."

The proteins produced by tumour suppressor genes, meanwhile, normally prevent cancerous growth. If the genes are damaged

or lost, cells are more likely to become cancerous. The *p53* and *Rb* tumour suppressors, for instance, are inactivated in most tumours⁸. But proteins themselves do not make good drugs — they are hard to administer, and tend to get broken down in the body — so scientists are now trying to block key proteins in the molecular signalling pathways given free rein when tumour suppressor genes are inactive.

Many genes associated with cancer interact with several signalling pathways. For example, *p53* promotes apoptosis and activates DNA repair — and probably also inhibits angiogenesis⁹. Many oncogenes encode enzymes called tyrosine kinases, which add a phosphate group to a protein in a molecular pathway as a means of propagating the signal. Again, this activity can operate in several different signalling pathways.

Initially, drug development lagged behind the explosion of studies of cancer genes and the pathways that they influence. "The flood of identified oncogenes was very

interesting," says Barbacid. "But we got to a point where it was just another oncogene, another kinase, another pathway — and it was definitely time to put the knowledge in the service of the patient."

When drug companies moved in, they were particularly interested in cancer genes encoding tyrosine kinases, because the pharmaceutical industry has decades of experience in finding small molecules to block specific enzyme targets. Dozens of tyrosine kinase inhibitors are now being tested in the clinic, and Gleevec is the first to make it through.

Toxic twist

Gleevec was originally designed as a specific inhibitor for the platelet-derived growth factor receptor, which acts as a tyrosine kinase and has been implicated in some cancers. But the compound was later found to interact with proteins produced by two other oncogenes, *BCR-ABL* and *c-KIT*. The former — a mutation caused by the fusion of sequences in the *ABL* and *BCR* genes — is the trigger for CML¹⁰. It causes a recognizable chromosomal defect called the Philadelphia chromosome, in which chromosome 22 is shorter than normal.

Initially, pharmacologists were worried that Gleevec would have dangerous side effects. They were particularly concerned about its interaction with *c-KIT*, which, when functioning normally, is involved in regulating the immune system. "The fact is that Gleevec should have been toxic," says Barbacid, "but it wasn't." No one understands exactly how, but the immune system seems to be able to switch on other pathways to compensate for the blocking of the tyrosine kinase encoded by *c-KIT*.

Since its approval by the FDA in May 2001, positive clinical results^{1,2} have led to Gleevec's licence being extended in February this year to a rare stomach cancer, gastrointestinal stromal tumour (GIST) — a disease that is caused by the mutation of *c-KIT* (ref. 11).

Because they are caused by single mutations, CML and GIST are 'easy' diseases. But tyrosine kinase inhibitors are also being tested in messier situations — in cancers where many mutations have accrued. Even Gleevec does not perform quite so well in these circumstances. In a later and frequently fatal stage of CML called blast crisis, where many mutations have appeared, only around two-thirds of patients respond to the treatment¹. Resistance to the drug also develops quite quickly as *BCR-ABL* mutated its way free of Gleevec's effects, or was overexpressed.

But at least one other tyrosine kinase inhibitor has performed well in early clinical trials against a notoriously difficult cancer. AstraZeneca's Iressa, directed against the tyrosine kinase activity of the EGF receptor, has excited oncologists by prompting a 10% response rate in patients with non-small-cell

The real deal in Madrid



Four years ago, Mariano Barbacid was made an offer he couldn't refuse: a chance to return to his native Spain to direct a new national cancer centre, with carte blanche to structure it however he saw fit.

In the early 1980s, while at the National Cancer Institute in Bethesda, Maryland, Barbacid led one of the teams that cloned the first human oncogene, *ras*. Later, as vice-president for oncology drug discovery at Bristol-Myers Squibb in Princeton, New Jersey, he absorbed the drug-industry culture. At the Spanish National Cancer Center, which moved into its new building in Madrid in February, Barbacid now aims to marry basic research with drug development.

The centre will conduct basic research into molecular and genetic oncology, but half of the 500 staff will work on applied projects in diagnostics and drug discovery. There will even be a medicinal chemistry programme to develop candidate drugs.

This is unusual for a cancer centre anywhere. It is

virtually impossible for those in the United States to find funding for medicinal chemistry, says Frank McCormick, director of the Comprehensive Cancer Center at the University of California, San Francisco. "We form collaborations with pharmaceutical or biotech companies."

According to Barbacid's philosophy, this may often be just as well. "I've known some very clever scientists who say idiotic things about drug development," he says. But given Barbacid's experience in industry, he is optimistic of making progress — partly in-house, partly through links with drug companies. "We will concentrate on one or two targets," says Barbacid. "We don't expect to compete with big companies, but we may be lucky."

The centre's structure is designed to court good fortune. Barbacid has a staff member located in each of nine hospitals in the Madrid area to ensure good contacts with clinicians, and a supply of samples for the centre's tumour bank — more than 2,000 have already been collected. He is developing strong core facilities, including housing for 100,000 mice, plus facilities for microarray manufacture and the structural analysis of drug targets.

In practice, Barbacid is largely restricted to hiring Spanish researchers. "Spain cannot compete for postdocs internationally," he concedes. Fortunately, there are plenty of talented young scientists around, who have trained abroad and survived on short-term contracts since their return.

What Barbacid does not have is assured continuity of funds. The government currently covers half of the operational costs, but this has to be agreed annually — and Spanish politics can be fickle. The support of Spain's popular king has helped him so far, as has Barbacid's go-getting personal style. But that approach has made a few enemies, as well as friends — so the pressure is on to succeed.

Alison Abbott, Madrid

► www.cnio.es/english

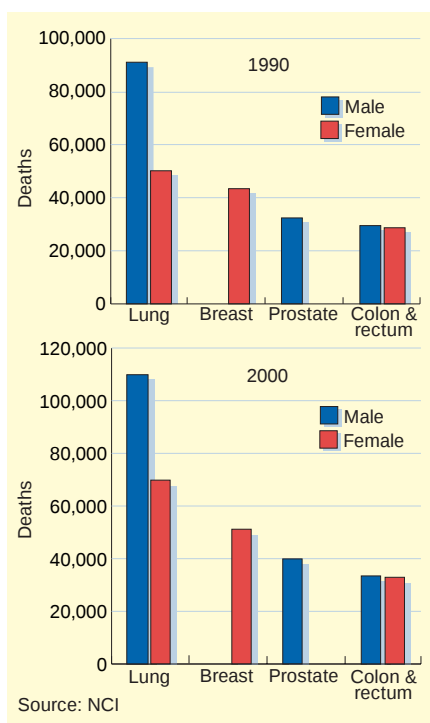
lung cancer who had failed to respond to other therapy. Although this might sound like a low response, the outlook for these patients is usually extremely bleak. "This is miracle-like," says Charles Sawyers of the Jonsson Comprehensive Cancer Center at the University of California, Los Angeles, who ran some of the Gleevec trials.

Road blocks

Other classes of drug have so far not performed as well, although efforts are continuing. A few years back, for instance, inhibiting angiogenesis by blocking the action of proteins such as vascular endothelial growth factor was thought to hold great promise. One front-page story in *The New York Times*, published in May 1998, infamously touted angiogenesis inhibitors as the long-sought cancer 'cure'. But clinical trials of the drugs have so far disappointed — cancer cells seem to find it easy to get around the blockage of one angiogenic pathway.

There is also less excitement now about the potential for telomerase inhibitors, mostly because of concerns about their toxicity to stem cells in bone marrow, which also require the enzyme. And attempts to disrupt metastasis — using drugs called matrix metalloproteinase inhibitors to block an enzyme used by cells to chew their way out of the extracellular matrix that usually keeps them in place — have so far disappointed in the clinic¹².

Given previous experience, are researchers getting too excited about Gleevec? No, argue enthusiasts for the molecular approach. First, Gleevec is performing well where it matters — in the clinic. Second, no one is pinning their hopes on one drug alone. Gleevec is merely the proof of principle that you can block a pathway that cancer cells depend on. In most cancers, it may be necessary to block pathways at several points, or even to target several



Slow progress: US mortality rates for the major cancer killers have changed little in 10 years.

pathways. So, repeating the success of Gleevec in other cancers may require cocktails of drugs — perhaps including some of those, such as the angiogenesis inhibitors, that have performed poorly in isolation.

"It is naive to think that in solid tumours we will get dramatic results by targeting one gene," says José Baselga, a clinical pharmacologist at the Vall d'Hebron University Hospital in Barcelona, who is organizing 58 clinical trials of tyrosine kinase inhibitors. "There will be many mutations, and so in future we can reckon on using combinations of drugs to hit many targets."

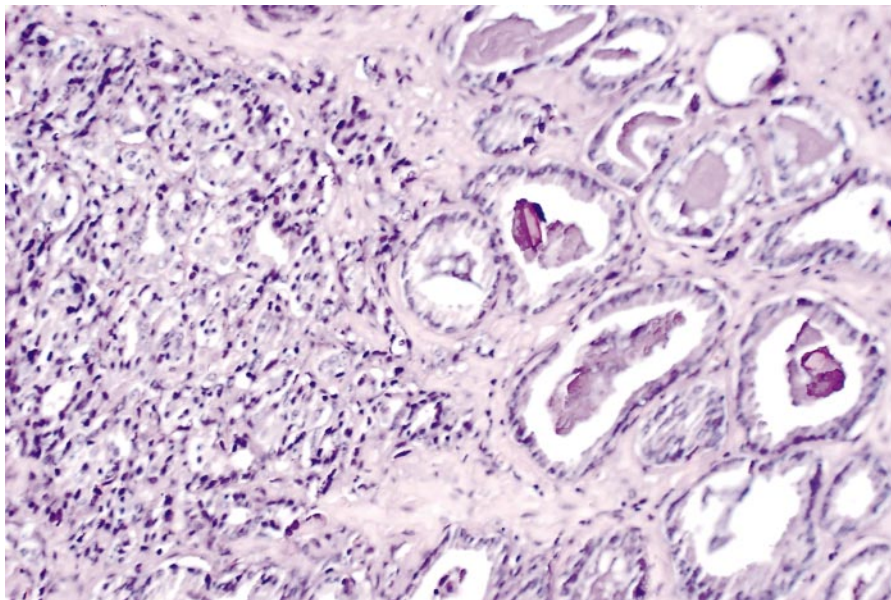
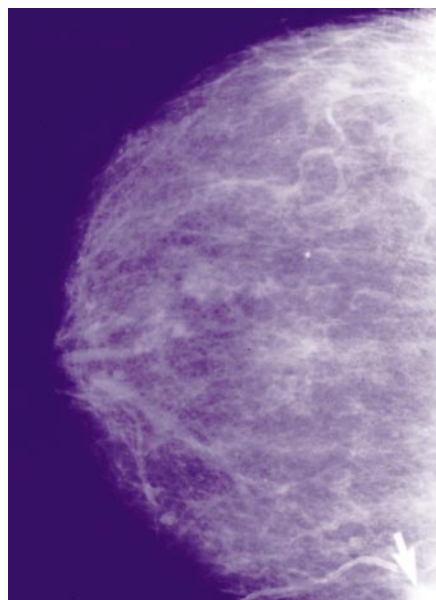
"This is what happened with AIDS treatments," agrees Edward Sausville, associate director of the NCI's developmental therapeutics programme. "Individual drugs did little on their own — but they worked very powerfully when they were put together."

Specific drug combinations will probably need to be tailored to particular tumour types — which is why efforts to profile individual tumours to find out what, exactly, has gone wrong at the molecular level form an important new front in the war on cancer. "The first step is to understand the molecular profiles of cancers — then we'll need to identify targets," says Robert Strausberg, head of the Cancer Genome Anatomy Project at the NCI.

Biological bull's-eyes

In the past few years, huge investments have been made in new genomic technologies to do just this. Strausberg's project is the largest of many attempts to build up databases storing information on gene and protein expression, gene mutation and silencing, and clinical information such as the rate of disease progression and response to drugs. According to Sausville, these projects will lead to the recognition that types of cancer currently viewed by pathologists as identical are in fact distinct molecular diseases. "The most important thing we will learn is how to classify cancers better," he says.

Two recent papers, using DNA microarrays to look at gene expression in a lymph-node cancer called diffuse large B-cell lymphoma¹³ and in breast cancer¹⁴ have illustrated the point: within each disease, cancers can be grouped into subsets with distinctive gene-expression profiles that correlate with how fast the disease progresses. Sawyers also suspects that the 10% response to Iressa seen in the lung cancer trials indicates that the tumours that responded had a



Fresh ammunition: could drugs targeted at specific molecular pathways provide a weapon against the 'big' killers such as breast (left) and prostate cancer?

distinct molecular profile. "My hunch is that the EGF signalling pathway is the driving force in 10% of lung cancers, and the other 90% were different diseases," he says.

Molecular profiling should also allow scientists to identify markers for early diagnosis of cancer — proteins that could be measured in blood, faeces, urine or even in shed skin. "Early diagnosis is very important because the cure rate is very high for early-stage disease and very low for late-stage disease," says Lee Hartwell, director of the Fred Hutchinson Cancer Research Center in Seattle, who last year shared the medicine Nobel for his work on the cell cycle of growth and division.

Bruce Ponder, head of the University of Cambridge's Department of Oncology, and co-director of a new cancer centre being established at the university by the Medical Research Council and the charity Cancer Research UK (see 'Two into one', below), hopes to find molecular indicators of

whether new drugs are working in clinical trials. "At the moment, tumour shrinkage is used as an endpoint and this is not only crude, but also delayed," he says.

Molecular profiling might also help clinicians to use currently available drugs more efficiently. "By comparing accumulating data on molecular profiles with outcomes of clinical treatment, we'll be able to identify, for example, which patients will not benefit from chemotherapy after surgery and spare them from useless, and very unpleasant, treatment," says Ponder.

Informed choice

In parallel with gathering data on tumours, some scientists are profiling cell lines used to study cancer in the lab. John Weinstein, a molecular pharmacologist at the NCI, is analysing gene and protein expression in the NCI's 60 standard cancer cell lines. Over the years, these cells have been used to test

more than 70,000 different drugs. "There is a mine of pharmacological information there which we will correlate with the changes in gene and protein expression, to help us work out what types of drug will work best in a cancer with a particular expression profile," Weinstein says.

Already, Weinstein's research has turned up interesting pharmacological insights. For example, the enzyme L-asparaginase is used to treat acute lymphoblastic leukaemia (ALL) because it destroys the amino acid L-asparagine in the blood. This works because ALL cells are unable to make their own supply of the amino acid, so once their access to it in the blood is blocked, they die. Weinstein has found that ovarian cancer cell lines have similar patterns of gene expression to ALL cells, which suggests that they might also be sensitive to L-asparaginase¹⁵.

Molecular profiling will yield masses of data to add to our existing understanding of the signalling pathways that influence cancer. But from this complexity, leading researchers are convinced that simple insights will emerge. "Cancer biology and treatment... will become a science with a conceptual structure and logical coherence that rivals that of chemistry or physics," Hanahan and Weinberg argued in their 2000 review³. They claimed that within two decades, cell biologists will have derived a complete integrated circuit of the cell's signalling pathways, allowing us to model how specific genetic perturbations cause cancer, and to predict how to correct the problem using drugs acting on key points in the circuit.

Other experts agree that marrying this 'systems biology' approach with weapons such as Gleevec holds great promise — not of defeating cancer within a few years, as Nixon once promised, but hopefully of seeing real progress in the ongoing war on cancer over the next decade or two. "History shows that therapy comes when you understand the system of the disease," says Bert Vogelstein, a leading cancer researcher at Johns Hopkins University in Baltimore, Maryland. ■

Alison Abbott is Nature's senior European correspondent.

1. Savage, D. G. & Antman, K. H. *N. Engl. J. Med.* **346**, 683–693 (2002).
2. Shawver, L. K., Slamon, D. & Ullrich, A. *Cancer Cell* **1**, 117–123 (2002).
3. Hanahan, D. & Weinberg, R. A. *Cell* **100**, 57–70 (2000).
4. Goldfarb, M., Shimizu, K., Peruchio, M. & Wigler, M. *Nature* **296**, 404–409 (1982).
5. Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474–478 (1982).
6. Santos, E., Tronick, S. R., Aaronson, S. A., Pulciani, S. & Barbacid, M. *Nature* **298**, 343–347 (1982).
7. Bos, J. L. *Cancer Res.* **49**, 4682–4689 (1989).
8. Sherr, C. J. *Nature Rev. Mol. Cell Biol.* **2**, 731–737 (2001).
9. Dameron, K. M., Volpert, O. V., Tainsky, M. A. & Bouck, N. *Cold Spring Harb. Symp. Quant. Biol.* **59**, 483–489 (1994).
10. Shivelman, E., Lifshitz, B., Gale, R. P. & Canaan, E. *Nature* **315**, 550–554 (1985).
11. Hirota, S. *et al. Science* **279**, 577–580 (1998).
12. Coussens, L. M., Fingleton, B. & Matrisian, L. M. *Science* **295**, 2387–2392 (2002).
13. Alizadeh, A. A. *et al. Nature* **403**, 503–511 (2000).
14. van't Veer, L. J. *et al. Nature* **415**, 530–536 (2002).
15. Scherf, U. *et al. Nature Genet.* **24**, 236–244 (2000).
16. Petricoin III, E. F. *et al. Lancet* **359**, 572–577 (2002).

Two into one

In Britain, the government plays second fiddle to the charitable sector in defining the agenda for cancer research. And in February, that became even more obvious with the merger of the two largest British cancer charities — the Cancer Research Campaign (CRC) and the Imperial Cancer Research Fund (ICRF) — to form Cancer Research UK. Officials argue that the merged organization, which boasts an annual budget of about £150 million (US\$214 million) and supports some 3,000 researchers, is better placed to meet the challenges of converting basic biological insights into effective new therapies. The ICRF, they point out, largely focused on fundamental research whereas the CRC carried out more clinical work.

"The merger has the potential to offer more streamlined drug discovery," says Linda Lashford, director of translational research at the new charity. "It should be much clearer to people who has the skills to move the process from one stage of the pipeline to the next." Access to equipment and resources should also improve, she predicts. For example, ICRF researchers will have easier access to a library of some 50,000 small molecules, used to screen against candidate drug targets, held at the CRC's laboratories in Sutton, near London.

Discussions about the merger were made public more than a year ago, but many details remain to be finalized — not least who will ultimately lead the new organization. ICRF chief Paul Nurse, who shared last year's medicine Nobel, and Gordon McVie, head of the CRC, have for now been made joint director-generals of the merged body. The post of interim chief executive has gone to an outsider, Andrew Miller, previously the vice-chancellor of Stirling University in Scotland and a former head of the European Molecular Biology Laboratory's outstation in Grenoble, France. Miller's role is to steer the merger through its critical early months. He is combining the financial and administrative systems, setting up a senior management structure — part of the task, he says,

"is to replace myself" — and trying to prevent the new organization from being dominated by the culture of one of the pre-merger charities.

One tricky issue is the contrasting funding methods used by the ICRF, which carried out most research in its own institutes, and the CRC, which spent the bulk of its money on independent groups in universities. "The intention is to maintain both types of funding but to bring them closer together in terms of judging quality criteria," Miller says.

Although most British cancer researchers seem to be happy with the merger, some sceptics have suggested that the new organization could homogenize research efforts and stifle innovative ideas. McVie disagrees, saying that a significant amount of the charity's budget will be set aside for "quick response" research funding. "That's exactly the kind of thing you could use to test an off-the-wall idea," he says.

David Adam, London

▶ www.cancerresearchuk.org



Joined forces: Cancer Research UK's Paul Nurse (left) and Gordon McVie.

Miller revealed new ways to study the origins of life

Science advances as one theory builds on another: Miller didn't just update Löb's work.

Sir — Hubert P. Yockey in Correspondence (*Nature* **415**, 833; 2002) claims that Stanley L. Miller's classical synthesis of amino acids and other organic compounds with electric discharges using possible prebiotic conditions (*Science* **117**, 528; 1953) was a mere repetition of several previous electrochemical syntheses performed by Walther Löb and others in the early twentieth century. This is not a fair assessment of the significance of Miller's experiment.

Although some of Löb's results may have some bearing on our understanding of prebiotic syntheses, part of the significance of Miller's experiment lies not only in the production of amino acids and other compounds, but in their synthesis under what was viewed at the time as plausible primitive Earth conditions. Few, if any, scientific ideas are the product of spontaneous thoughts — most theories, experiments and interpretations have been preceded by many others, and the same is true of Miller's experiment. Even if one disagrees with the assumptions underlying the simulation by Miller and Harold Urey of the primitive Earth, it deserves recognition not only because of its intrinsic merits, but also because it opened new avenues of empirical research into the origin of life.

Löb did indeed report the synthesis of glycine by exposing wet formamide to a silent discharge (*Ber. Dtsch Chem. Ges.* **46**, 684; 1913). He suggested that because of either the ultraviolet light or the electrical field generated by the silent discharge, formamide is first converted to oxamic acid, which in turn is reduced to glycine. He also claimed that glycine is produced when wet carbon monoxide and ammonia are subjected to the silent discharge; he proposed formamide as the intermediate in this synthesis. Löb theorized that glycine might also be produced from wet carbon dioxide and ammonia in a pathway wherein formamide was again the intermediate, but he did not demonstrate this directly.

Although Löb apparently did produce glycine from formamide, this cannot be considered a prebiotic reaction because formamide would not have been present on the primitive Earth in any significant concentrations. It is also possible that the wet carbon monoxide and ammonia led to HCN synthesis, which would have produced glycine on polymerization and hydrolysis.

From a careful reading of Löb's 1913 paper (in early twentieth-century German!) it is clear that his motivation for doing the experiment was to try to

understand the assimilation of carbon dioxide and nitrogen in plants. There is no indication that he had any interest in the question of how life began on Earth, or in the synthesis of organic compounds under possible prebiotic conditions. Neither Aleksander Oparin, J. B. S. Haldane nor Urey made any mention of Löb's work, which given Oparin's extensive review of early relevant literature suggests it was considered unimportant. To the best of

our knowledge, Löb's work was first discussed within the context of prebiotic chemistry by Miller (*J. Am. Chem. Soc.* **77**, 2351; 1955).

Jeffrey L. Bada* and **Antonio Lazcano†**

**Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California 92093-0212, USA*

†Facultad de Ciencias, UNAM/Apdo. Postal 70-407, Ciudad Universitaria, 04510 Mexico D. F., Mexico

Communication should not be left to scientists

Sir — The time has come for the professionalization of science communicators. Not only are the number of degree programmes in science communication growing, but rudiments of a professional code of conduct have now been published. Although there is nothing new about the advice given in *Guidelines on Science and Health Communication*, issued in November 2001 by the Social Issues Research Centre in partnership with the Royal Society and the Royal Institution (see www.sirc.org/publik/revised_guidelines.shtml), what is striking is that it is directed at science communicators by people who are not themselves full-time science communicators.

This pattern is familiar from the history of such 'mediating' professions as nursing and librarianship. Members of a senior profession typically try to formulate the aims of the newer profession in terms of their own interests. Physicians have tried to set nursing codes, academics librarian codes, and now scientists are trying to set codes for science communication.

At the same time, however, the mediating professions have often provided their own visions, which in significant respects cut against those of the senior professions. For example, Florence Nightingale did not want nurses to make it easier for doctors to administer medicine, but to provide a total healing environment that would eventually prevent the need for doctors. Similarly, Melvil Dewey's cataloguing system envisaged librarians, not as dutiful retrievers of books, but as the last encyclopaedists in a world where knowledge was quickly fragmenting.

To be sure, nursing and librarianship have not been completely successful in their professionalization. Nevertheless, their histories provide valuable lessons for science communicators. The public profile

of nurses and librarians was raised as physicians and academics, respectively, became increasingly specialized. This led ambitious members of the mediating profession to occupy the role previously filled by the 'general practitioner' of the senior profession: someone sufficiently familiar both with a broad range of specialist knowledge and with clients' particular needs to make the knowledge meaningful for them.

If the history of the mediating professions is a guide, the idea of science communicators as general practitioners of science may well generate conflict with professional scientists. Whenever someone claims that Richard Dawkins's *The Selfish Gene* has done more for biology than most front-line research in the field, we get a taste of those potential battles.

Nevertheless, by involving the general public, science communication is uniquely placed to make good the idea that science is truly universal knowledge. The current guidelines reduce communication to the one-way transmission of accurate information. More ambitious professional guidelines would require a feedback element in any form of science communication, whether it be as letters to the editor, Internet chat rooms or the interactive exhibits now routine in good museums.

The current guidelines pitch the aims of science communication too low. The field should aspire not simply to make people trust science but to make them feel part of it. This is more than encouraging people to become scientists. Science communication will truly come of age when people regard participating in a scientific experiment as a civic duty on par with jury duty. We are far from that day, but it is not too early for full-time science communicators to convene formally to draft some professional codes of conduct.

Steve Fuller

Department of Sociology, University of Warwick, Coventry CV4 7AL, UK

Making light work of computing

A lucid account of optical information-processing technology.

Mind at Light Speed: A New Kind of Intelligence

by David D. Nolte

Free Press: 2001. 320 pp. \$26

Ian Walmsley

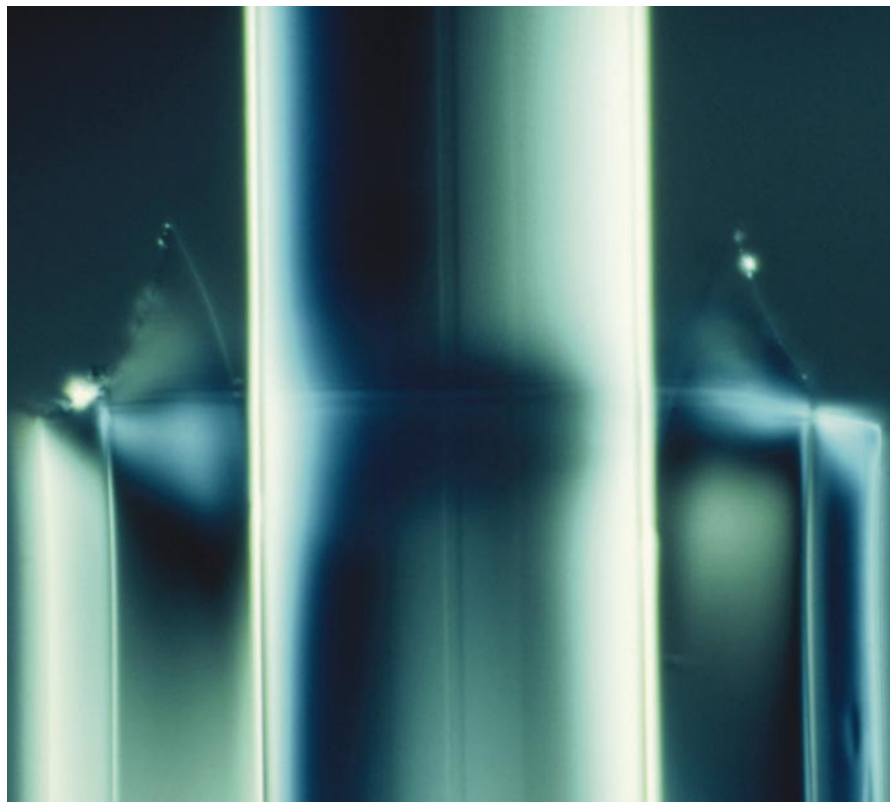
As a professional scientist, it's never easy to explain to your mother what you do for a living unless she happens to be one too. So most of us find ourselves in the situation described by David Mermin in the preface to his book *Boojums All the Way Through* (Cambridge Univ. Press, 1990). The response to any explanation is invariably: "It's very impressive — I couldn't understand a single word."

Nolte's book is the antidote to such a dilemma. It's the sort of thing one could hand to a relative, even one not versed in university-level science, and expect them to be able to understand something of the field of optical information processing. The core of the book describes a hierarchy of optical networks: the optoelectronic, the photonic and the quantum. Nolte uses this structure as a hook on which to hang lucid explanations of the relevant technology for each, and its limitations.

To explain how information is processed, the book begins with some physiology and neuroscience, and even semiotics. Some of this is relevant to modern information technology: neuronal processing of sensory information has direct descendants in computational pattern recognition, for example. But it seems a stretch to claim that the rate of sensory information input is really a bottleneck to intelligence.

Nolte then goes on to explain how optics can dramatically increase the information capacity of a communications system, by combining the inherently large potential bandwidth of optical-frequency waves and the near inability of light to interact with itself in glass. Controlling (that is, sorting and routing) the optical data packets along the strands of glass that form the backbone of the current optical network requires particles that are not so benign. In optoelectronic systems these are electrons, but Nolte makes a convincing case that light can act as both Mercury and Apollo — messenger and diviner — through the development of new nonlinear optical materials.

The intensity-dependent response of optical detectors makes possible the final ingredient in classical optical information processing: holography. The description of this concept and its use in information storage is a delightful part of the book, and the back-of-the-envelope calculations of



Signal strength: fibre-optic technology has allowed vast quantities of data to be stored and transmitted.

the storage capacity of a holographic memory are something I would strongly recommend to most optics students. Nolte shows elegantly the idea of optimal encoding that leads to the massive parallelism that optics can provide.

The final section of the book concerns the much more futuristic technology based on information processing using quantum systems. The potential of such systems for secure communications and faster computations is harder to explain than the classical physics associated with the current and next generations of optical technologies. Some parts of this section are insightful. For instance, the notion that information does not exist in the quantum system until it is observed implies that the 'parallel processing advantage' of quantum interference and entanglement has a more subtle meaning than one might naively imagine.

But other parts of the book are rather misleading. For instance, the explanation of the Einstein-Podolsky-Rosen paradox almost misses the point entirely. Correlated particles that exhibit the same sort of non-locality between distant measurements, as Nolte describes, exist in the classical world. It is the ability simultaneously to ascribe both a definite position and momentum to the

second of two non-interacting particles, in apparent contravention of Heisenberg's uncertainty principle, that led Einstein, Podolsky and Rosen to posit that quantum mechanics did not satisfy the criteria expected of a physical theory.

The thesis that Nolte develops sporadically throughout the book — that optics is the route to intelligent machines — is perhaps overstated. But the fervour with which he describes it conveys the excitement engendered by working in cutting-edge science. Indeed, he seems as enchanted with his optical technology as the smitten Pygmalion with his creation Galatea.

If one overlooks the speculative excesses of the first and final chapters, Nolte has a remarkable ability to bring out the essential scientific features in a way that is both thought-provoking and pedagogical. He provides a fairly complete picture for the student and interested amateur of why the technology works the way it does, describes the roadblocks to improving system performance, and discusses the effects on telecommunications and data processing. You can always skip the teleology. ■

Ian Walmsley is at the Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

Fixed ideas take root

Nodulation in Legumes

by Janet I. Sprent

Royal Botanical Gardens: 2001. 146 pp. £27

Sharon R. Long

The formation of nitrogen-fixing nodules on the roots of legumes (peas, beans and their relatives) is the basis of crop rotation, a mainstay of sustainable agricultural practice. These nodules are the result of an association between the legumes and *Rhizobium* and allied bacteria. This symbiosis is a wonder of evolution, and has for years been a target for molecular, genetic, physiological and population-level studies.

Many questions have been addressed through the study of model systems of bacteria and legumes. For example, how do bacteria and hosts recognize each other? How do they coordinate action so that the plant produces a completely new organ, the root nodule, which the bacteria invade and inhabit? Why do symbiotic bacteria express their enzymatic apparatus for nitrogen fixation only after developing into specialized forms within the nodule? And why do they fail to assimilate their own fixed nitrogen, instead exporting it to be assimilated by the plant?

And the mystery remains: why legumes? This question refracts into a spectrum of issues related to specificity and adaptation. Why are different legume species or genera nodulated by different (sometimes multiple) bacteria? Why do some legumes appear not to support nodulation? We assume that the plant gains some physiological advantage from the presence of the symbiosis, but how robust is this assumption? Do we know enough about the physiological constraints that limit plant productivity for diverse legumes?

The scope of this second set of questions is as large as the legume family itself (it is the third-largest family of flowering plants) and

as wide as the extent of green Earth, for legumes inhabit the tropics and the Arctic, deserts and rainforests, beaches and mountains. Janet Sprent's *Nodulation in Legumes* is especially valuable because it reminds us that model systems represent only a small part of a big story: how legumes and their rhizobial symbionts work together in the real world.

Early chapters present a concise overview of diversity, physiology and mechanism. Legumes exist in widely varied environments, and their growth may be subject to varied constraints. During nodule development, bacteria and plant communicate through a set of signals, and new structures and new gene expression characterize both early and late stages of nodulation. The physiology of the nitrogen-fixing state involves careful control of (and often regulation by) the level and the flows of oxygen, nitrogen, carbohydrates or organic acids. These chapters are readable and compact, and do not generally refer to the fast-expanding primary literature on mechanisms of nodulation. Because this research area is dynamic, with conclusions and generalizations constantly being re-examined, students interested in going further should consult more specialized reviews.

A series of further chapters survey nodulation in each of three legume subfamilies: Caesalpinoideae, Mimosoideae and Papilionoideae. These are especially valuable given the direction in which laboratory science is heading: complete genome-sequencing analyses have recently shown surprising diversity at the whole-genome level of various rhizobia. In parallel, plant genetics and genomics will soon make it possible to look in diverse legumes for symbiosis genes encoding receptors and response pathways. For the next generation of legume researchers, Sprent's book will provide a platform for the expansion of study from a few model systems to a wide world of bacteria and legume hosts.

Questions of evolution and diversity are addressed in the final chapter, using data from the fossil

record, morphological analysis and molecular taxonomy, for example. It is extremely valuable to have so much work summarized in one place. Sprent speculates that ancestral legumes had to acquire two features in order to support symbiosis: the suppression of host defences and the pathways to build the nodule structure — a structure that shows considerable variation among members of the family. To these I would suggest a variation: legumes must have the active cellular machinery to support infection, whether or not the plant sustains enough cell divisions to form a nodule. Which prompts me to wonder whether the ancestral events for modern-day nodulation were based in infection, rather than morphogenesis *per se*, and whether these events might be found in extant legumes that show no visibly evident formation of nodules.

A second printing of the book would provide the opportunity to correct some minor typographical errors that turn up in the table of genera (for example, *Daniellia* is spelled in two ways, so it looks like two genera in separate parts of the book). But this is a minor issue for a valuable book that *Rhizobium*–legume researchers and plant-evolution scholars alike should have and should study. As well as reminding us how broad and diverse is the world of symbiosis, and beyond the valuable summaries of physiological and ecological studies, the author conveys a lifetime perspective that places these elements into a framework of knowns and unknowns, of questions central and ancillary. And through it all, every reader of this book will come away feeling the author's love for her subject. The caption to Figure 1.2 begins "Nodule hunting!" That exclamation mark says it all. ■

Sharon R. Long is in the Department of Biological Sciences, and is dean of the School of Humanities and Sciences, Stanford University, Stanford, California 94305, USA.

Lightning on the veld

Schonland: Scientist and Soldier

by Brian Austin

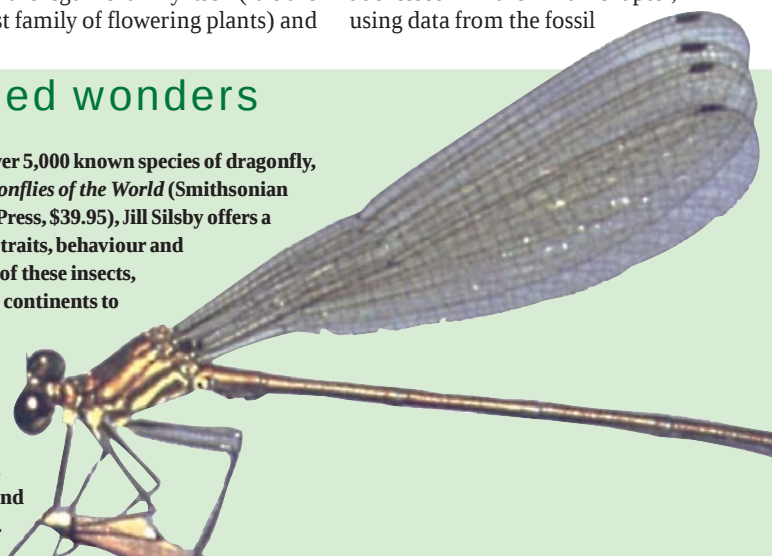
Institute of Physics Publishing: 2001. 639 pp. £49, \$73.

John D. Hey

The life of Sir Basil Schonland CBE FRS (1896–1972) spanned a crucial period of transition in geopolitical, military and scientific affairs — from the decline of empires to the Cold War and the nuclear-based arms race, and from classical physics to the physics of relativity and the quantum that underpins the technology of the space age.

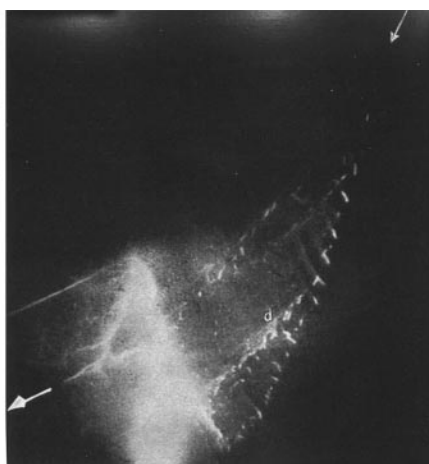
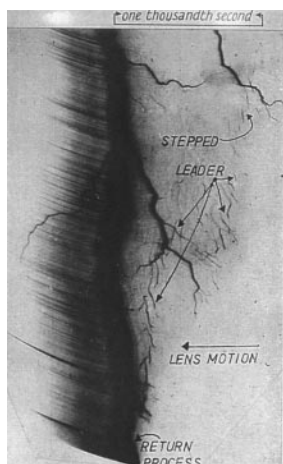
Winged wonders

There are over 5,000 known species of dragonfly, and in *Dragonflies of the World* (Smithsonian Institution Press, \$39.95), Jill Silsby offers a guide to the traits, behaviour and appearance of these insects, crossing the continents to include specimens such as *Megalestes kurahashii*, shown here, which is found in Thailand.





With his photographs of lightning taken on the South African veld (see inset), Basil Schonland established a reputation as a world leader in lightning research.



Schonland contributed to both military and scientific developments through a mastery of physics, learned initially in the university town of Grahamstown in the Cape, South Africa, but perfected at the Cavendish Laboratory during the 'golden age' of Sir Ernest Rutherford.

A product of the English type of public school, young Basil's career combined brilliant scholastic triumphs with sturdy character development. He acquired from early youth a belief in the ideals of service to King and Country, and also qualities of leadership, which guided this intellectual brilliance to the benefit of all who later came under his authority.

Austin develops the multi-faceted tale of his achievements in vivid detail. We encounter Schonland as scholar and student, officer and communications specialist in the Royal Engineers, wireless research officer and chief instructor of the British Expeditionary Force, recipient of the OBE (1919), and postgraduate student at Cambridge (1919–1922). Here he grappled with an intriguing scientific puzzle: why the scattering of β -particles (energetic electrons) by thin metal sheets appeared to contradict the Rutherford atomic model. Meticulous experimental work was supported by the theoretical analysis of Charles G. Darwin.

Awarded his PhD in 1924, Schonland

which the huge natural laboratory of the open veld offered untapped opportunities. Schonland established a reputation as a world leader in lightning research and in 1938 was elected FRS; he was by then director of the newly established Bernard Price Institute (BPI) of Geophysical Research in Johannesburg.

Considerable experience in radio work by Schonland and his team at the BPI by 1939 and three days' tuition in radar by Ernest Marsden, fortuitously en route to New Zealand, facilitated the successful construction of portable radar sets in Johannesburg from locally available components. These proved invaluable to the war effort in East Africa and the Middle East, when British sets were in short supply, and provided the stepping-stone for Schonland's re-entry into active military service. His major contribution here was the practical implementation of operations research, in his capacity as scientific adviser to Field Marshal Montgomery, an appointment facilitated

by impeccable military credentials.

Brian Austin, an expert on research into 'wireless', radio and radar development, succeeds well in covering these aspects of his vast canvas. The text is well referenced and sources include records of interviews with the Schonland family, and former colleagues and students, who themselves played important parts in the story. The book also provides unique insight into the Cavendish at the zenith of the 'old' (Bohr–Sommerfeld) quantum theory. The text and group photographs, including several present and future Nobel laureates, remind us vividly that this period (1919–1928) was a golden age for the Cavendish. Two Nobel laureates (Edward Appleton and Charles Wilson) also played significant roles in Schonland's lightning research, while two others, close contemporaries (Patrick Blackett and John Cockroft), were vitally important to his military activities in the Second World War, to some of his geophysical investigations, to his successful planning and founding of the South African CSIR (1945–1950), and, especially, to his entry (1954) into the troubled fields of nuclear fission and fusion. Contemporary researchers in controlled thermonuclear fusion are provided with a salutary lesson on the dangers of prematurely publicized over-optimism and resource-consuming mammoth experiments, both problems that Schonland encountered in 1958.

The picture of Schonland the man emerges slowly, and is revealed only in the closing pages of the book. Clearly, public service of the highest order, without self-interest, remained pre-eminent throughout, even when the call of duty meant partly sacrificing a promising research career. Thus, ideals acquired in youth remained his lodestone right up to the point at which Schonland, "aged beyond his years", retired from management of the fractious world of large-scale British nuclear science (1954–1961).

Minor blemishes aside, this historical biography is a fine tribute to a legendary and inspiring figure in South African science. It may also serve to inform the reader of different aspects of a country, long under its own thundercloud, and eschewed, even in scientific circles, like the lepers of old — yet, as Thomas Hardy wrote of Egdon Heath, "perfectly accordant with man's nature, but like man, slighted and enduring". ■

John D. Hey is at the School of Pure and Applied Physics, University of Natal, Durban 4041, South Africa.

Edwin Black *Nature* published a correction (*Nature* **415**, 370; 2002) citing two errors in a review by Professor Howard Segal of the book *IBM and The Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation* by Edwin Black (*Nature* **411**, 993–994; 2001). In addition, Professor Segal, in his review, asserted that the book's author exaggerated the importance of punch-card machines. *Nature* has been made aware that the reviewer has retracted this assertion, stating: "In fact, I know of no evidence or documentation that the book included any exaggeration at all."

A risky business

Marcello Lotti and Pierluigi Nicotera

Several key advances in biology and medicine in the past were brought about by studies of poisons. The elucidation of the mechanism of carbon monoxide toxicity by Claude Bernard, which led to the understanding of the function of haemoglobin, is a classic example. However, Jean-Pierre Changeux and colleagues, who used α -bungarotoxin to purify acetylcholine receptors, and William Catterall, who isolated sodium channels using scorpion toxin, probably did not consider themselves to be toxicologists. Has modern toxicology actually provided new fundamental concepts? Surprisingly, with a few notable exceptions, it has not — it is nowadays regarded as an applied science that is devoted to minimizing environmental health risks posed by chemicals, mainly through risk assessment.

At the same time, there is an emerging crisis of confidence in toxicology as an applied science that can effectively predict risk, as illustrated by the debate about servicemen exposed to depleted uranium from weapons during the Kosovo conflict in 1999. Although there is no evidence for radiological or chemical carcinogenic risk at any conceivable level of exposure, the wide perception of this issue has been very different. Predictions based solely upon epidemiological projections without solid scientific bases are often misleading. Examples include the debates over genetically modified foods, dioxins, measles vaccinations, and prion diseases in

cattle and sheep. In the case of prion diseases, there is a remarkably uncertain and contradictory range of theoretical predictions for the size of any future epidemic of variant Creutzfeldt–Jakob disease in humans.

There are surely various reasons for this failure of trust, but we wish to discuss one in particular. Perhaps because of the immense scope of research into the mechanisms by which individual compounds act, basic research has, over the past two decades, become irrelevant to many toxicologists. A discipline that mostly depends on others for fresh fundamental knowledge, and is slow in acquiring it, will also be slow in its progress and weak in its conclusions. Prejudice, ideology and irrationality will undoubtedly grow. For instance, few among the public appreciate the fact that hazard and risk are different concepts. Hazard defines the potential of a compound to cause harm and is therefore associated with virtually any molecule, whereas a risk of adverse health effects relates to the level of exposure and to individual susceptibility to that molecule. Such misunderstanding may account for the generous public funds that have been allocated to the study of dioxin toxicity, despite the lack of evidence for effects on human health at current environmental exposures.

The exciting research opportunities that have arisen from recent spectacular developments in biology mean that we may finally have something to nail down and test. It is becoming increasingly clear that many forms of toxicity involve a handful of evolutionarily conserved responses to injury — for example, the heat-shock response or the mechanism of apoptosis. Understanding stereotypical reactions may help to identify the risks posed by a wide range of products, including genetically modified crops, new foods, chemicals and waste materials.

How can toxicological research remain a central discipline for risk assessment? To take cancer as an example, the molecular classification of some forms of the disease by monitoring gene expression suggests a general strategy to assess the risk associated with exposure to potential carcinogens. Tumour subtypes induced in animals by chemicals could be identified by developing algorithms to cluster cancers according to gene expression, assessing the significance of such aggregations and comparing them to those found in common human cancers. Thus, the rationale for current mutagenicity and carcinogenicity tests could be confirmed or refuted.

Tumour-cell genomes are invariably altered at multiple sites, and progressive genetic instability is closely associated with many cancers. Assessment of large gene

Toxicology

Toxicology research should urgently appraise its performance and join mainstream biomedical science.

rearrangements or genetic instabilities is likely to be very significant in these cancers. Comparisons of gene expression at different levels of carcinogen exposure should clarify the role of exposure for both cancer-associated and non-cancer-linked effects. Ultimately, comparing these patterns with those of common cancers could help to define whether animal carcinogens also pose a hazard to humans.

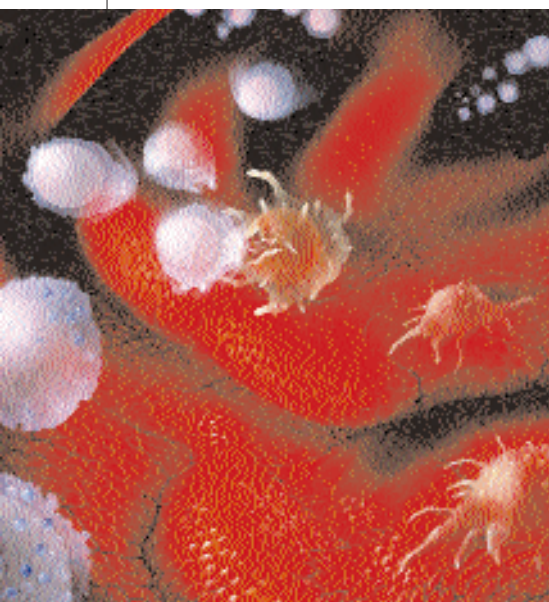
In a wider context, it seems possible to conceive a 'matrix' approach to the prediction of health risks, using information derived from commonality between disease states that are not induced by chemicals and those elicited by toxicants, and the differences between exposed and unexposed individuals.

Toxicology is being shaped by worldwide political agendas, triggered by the public's desire for swift and precautionary solutions to the possible health effects of environmental chemicals. The resulting feedback loop has impoverished the discipline, because its growth has largely been driven by the demand for protocols for regulatory actions. We believe that incorporating toxicology into the mainstream of fundamental biomedical research, keeping it less directly applied to issues of social and political concern, will contribute to a climate in which scientific knowledge is more 'socially robust', and its practitioners will accordingly be more trusted by the public. The assessment of toxicological risks ploughs a difficult furrow between scientific uncertainty and social responsibility. However, the approach we envision will hone the rules of hazard identification to deeper and more fundamental principles, while contributing to the development of basic biomedical knowledge. ■

Marcello Lotti is in the School of Medicine, University of Padua, via Giustiniani 2, 35128 Padova, Italy. Pierluigi Nicotera is in the MRC Toxicology Unit, Hodgkin Building, University of Leicester, Lancaster Road, Leicester LE1 9HN, UK.

FURTHER READING

Berry, C. *Trends Pharmacol. Sci.* **22**, 277–280 (2001).
Golub, T. R. *et al. Science* **286**, 531–537 (1999).
Hoeijmakers, J. H. J. *Nature* **411**, 366–374 (2001).
Smith, L. L. *Trends Pharmacol. Sci.* **22**, 281–285 (2001).
Gibbons, M. *Nature* **402**, C81–C84 (1999).
National Research Council. *Science and Judgement in Risk Assessment* (Nat'l Acad. Press, Washington DC, 1994).



Lymphocytes (pink) attack cancer cells (dull red). How should toxicologists best tackle the problem of assessing the risks of carcinogen exposure?

Danger — misfolding proteins

R. John Ellis & Teresa J. T. Pinheiro

Protein folding is vital to living organisms because it adds functional flesh to the bare bones of genes. But errors in this process generate misfolded structures that can be lethal.

The functions of most proteins depend upon the precise three-dimensional structures of their mature, folded forms. Take, for example, the most toxic substances known — proteins such as ricin, found in the seeds of the castor oil plant. The entry of a single molecule of ricin into an animal cell is in principle enough to kill that cell, because ricin is an enzyme that inactivates ribosomes, the cell's protein-synthesizing machines, one after another. The enzymatic activities of such toxins depend on their mature three-dimensional forms, but two papers in this issue^{1,2} suggest that some proteins pose a danger that is independent of either their mature structure or their activity.

On page 507, Bucciantini *et al.*¹ show that when two normally harmless proteins are each allowed to aggregate into fibrils *in vitro*, structures that form early in the aggregation process are highly toxic to cells, whereas the fibrils themselves are non-toxic. In similar vein, Walsh *et al.*² report on page 535 that early protein aggregates secreted from cultured cells expressing a mutant gene from a patient with Alzheimer's disease impair neuronal function in rat brains. The implication is that intermediates in protein folding can be dangerous in themselves, quite apart from any toxicity shown by mature folded proteins.

Most normal proteins are compact structures whose surface features determine their functions by presenting binding sites that are specific for other molecules. Such structures are produced by the process of protein folding, during which extended polypeptide chains, synthesized by ribosomes, collapse into more globular states in a manner determined by the amino-acid sequence of each chain. This collapse occurs because amino acids bearing hydrophobic side groups tend to huddle together inside the folded structure and thus avoid the aqueous environment. Correctly folded proteins are usually either soluble or associated with cell membranes. But the process of protein folding is prone to errors that generate aggregates of varying size, some of which are large enough to be insoluble. Such aggregates do not have the functions seen in correctly folded proteins, and are associated with some of the most distressing human diseases.

Aggregates arise because some intermediate structures that are formed during the folding process briefly expose hydrophobic

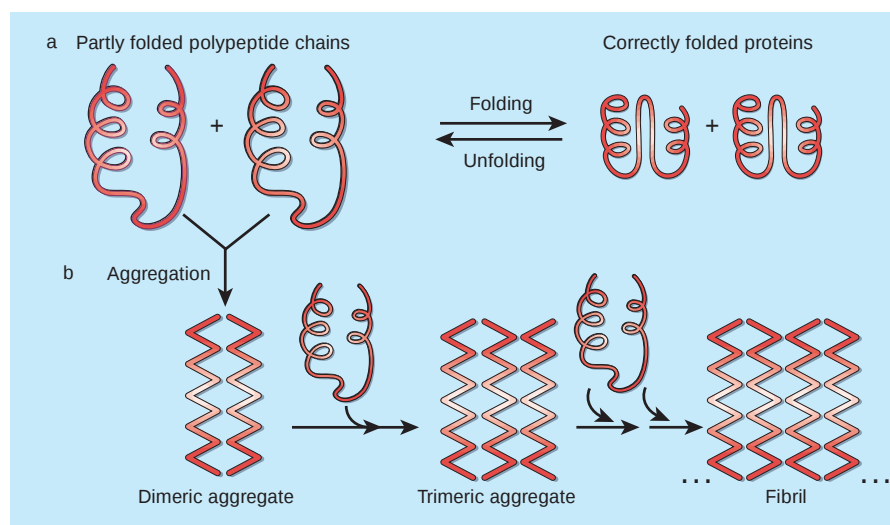


Figure 1 The problem of protein aggregation. a, Partly folded polypeptide chains, released from ribosomes (the protein-synthesizing machines), normally collapse into correctly folded, functional proteins. b, However, partly folded polypeptides may sometimes associate with similar chains to form aggregates. Partial unfolding of correctly folded proteins may also lead to aggregation. Aggregates vary in size from soluble dimers and trimers up to insoluble fibrillar structures. Unlike most correctly folded proteins, both soluble and insoluble aggregates can be toxic to cells through unknown mechanisms, as exemplified in the new papers^{1,2}.

regions on their surfaces, and these regions may bind to similar surfaces in nearby folding molecules instead of becoming buried inside the final structure (Fig. 1). The problem is made much worse by the high concentrations of macromolecules, including proteins, that are present in cells³, but is normally kept under control by proteins called molecular chaperones, which shield the exposed hydrophobic surfaces from one another⁴. Despite this control, several human neurodegenerative disorders such as Alzheimer's disease, and 'prion' diseases such as Creutzfeldt–Jakob disease, are associated with the occurrence of protein aggregates called amyloid fibrils, or plaques, as are protein-deposition diseases affecting other parts of the body such as the heart and liver. How these aggregates form, and how they cause disease, are topics of intense study⁵.

An intriguing property of amyloid fibrils that are associated with human diseases is that they can be formed from at least 20 unrelated proteins, but nevertheless show common structural features, such as a central core of 'β-sheets' (a common type of structural element) and the ability to bind dyes such as Congo red and thioflavin T. This commonality spawned the idea that

aggregation into amyloid fibrils is not specific to certain amino-acid sequences, but is a generic feature of all polypeptide chains under appropriate conditions, because β-sheet formation involves interactions between the atoms of main-chain amino acids that can occur in all proteins⁶. In support of this idea, proteins not known to be involved in disease, such as myoglobin⁷, can be induced to form amyloid fibrils *in vitro* when they are incubated under conditions that partly unfold their normal structure.

Bucciantini *et al.*¹ now find that when a structural region, or domain, from the bacterial regulatory protein HypF is incubated at pH 5.5 in the presence of the unfolding agent trifluoroethanol, aggregates form within a few minutes. The aggregates exhibit β-sheet structures and bind thioflavin T, but look amorphous when examined by electron microscopy, with no sign of fibrils. After 48 hours, short 'protofibrils' with disordered ends are seen, and after 20 days these are replaced by long, unbranched structures characteristic of mature amyloid fibrils.

Toxicity tests using cultured rat and mouse cells show that preparations of the amorphous aggregates or protofibrils kill these cells, as judged by the loss of their

ability to exclude the dye trypan blue, but preparations containing only mature fibrils do not¹. Similar experiments with a domain of a non-toxic bovine kinase enzyme also indicate that early *in vitro* aggregates kill the cells, unlike mature fibrils formed from the same part of the protein.

The second new paper² looks at the role of protein aggregates in Alzheimer's disease. Here, two types of aggregate are associated with brain damage — fibrillar tangles of the tau protein inside neurons, and amyloid fibrils of another protein outside. The amyloid fibrils contain many copies of the 40–42-amino-acid amyloid- β (A β) peptide. This is a fragment of a much larger protein of unknown function that is attached to the plasma membrane. The amyloid fibrils are toxic to cultured neurons, but whether they are the main cause of neuronal damage in Alzheimer's patients is unclear. The density of amyloid plaques in the brains of these patients is only weakly correlated with the severity of dementia. Moreover, previous work in several laboratories has shown that soluble forms of the A β peptide damage neuronal junctions (synapses)³.

Walsh *et al.*² now report that soluble dimers and trimers of the A β peptide, secreted from cultured cells, impair synaptic function when injected into rat brains, but that A β monomers, protofibrils and fibrils do not. The small A β aggregates specifically impair the ability of synapses to respond more strongly after receiving a rapid succession of stimulatory electric shocks. This property is termed long-term potentiation and is thought to be a key element in memory and learning. Encouragingly, drugs that inhibit the enzymes that produce the A β peptide reduce the formation of these dimers and trimers in cultured cells, hinting at potential treatments for Alzheimer's disease⁹.

Together, the results of Bucciantini *et al.*¹ and Walsh *et al.*² suggest that damage to cells can be caused by misfolded intermediates generated during the production of amyloid fibrils, whether or not the fibrils — or the normal proteins from which they are derived — are also toxic. The toxicity of these early aggregates depends upon some as-yet-undefined structural features, and not upon their amino-acid sequence. It is possible that some diseases not associated with the accumulation of amyloid fibrils may be caused by the sporadic production of this type of aggregate by mistakes that occur during the folding of other proteins.

These ideas highlight the importance of understanding the cellular mechanisms that combat these potentially lethal mistakes, and identifying the precise cellular targets of aggregates that escape such surveillance. The high variability in the age of disease onset suggests that aggregate production is triggered by the random, chance escape of a folding polypeptide chain from the

chaperone control machinery. Moreover, in the case of prion diseases, some misfolded chains are 'infectious' — they convert normal chains into the same misfolded structures.

One way forward may lie in the report¹⁰ that certain strains of *Escherichia coli* produce amyloid fibrils, using them to adhere to surfaces. This discovery suggests another approach to studying the mechanism of aggregate production — we can take advantage of how easy it is to genetically and biochemically manipulate this bacterium. ■

R. John Ellis and Teresa J. T. Pinheiro are in the

Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK.
e-mail: jellis@bio.warwick.ac.uk

1. Bucciantini, M. *et al.* *Nature* **416**, 507–511 (2002).
2. Walsh, D. M. *et al.* *Nature* **416**, 535–539 (2002).
3. Minton, A. P. *Curr. Opin. Struct. Biol.* **10**, 34–39 (2000).
4. Ellis, R. J. *Trends Biochem. Sci.* **26**, 597–604 (2001).
5. Dobson, C. M., Ellis, R. J. & Fersht, A. R. (eds) *Phil. Trans. R. Soc. B* **356**, 127–227 (2001).
6. Dobson, C. M. *Trends Biochem. Sci.* **24**, 329–332 (1999).
7. Fandrich, M., Fletcher, M. A. & Dobson, C. M. *Nature* **410**, 165–166 (2001).
8. Kim, J. H. *et al.* *J. Neurosci.* **21**, 1327–1333 (2001).
9. Selkoe, D. J. *Physiol. Rev.* **81**, 742–761 (2001).
10. Chapman, M. R. *et al.* *Science* **295**, 851–855 (2002).

Astrophysics

The evidence in the afterglow

Herman L. Marshall

The X-ray spectrum of an afterglow from a γ -ray burst reveals a smoking gun. It links the γ -rays to the expanding fireball that occurs after a supernova explosion.

Our understanding of γ -ray bursts (GRBs) becomes clearer with the new observations by Reeves *et al.*¹ on page 512 of this issue. GRBs are thought to result from the collapse of an extremely massive star to a black hole (prompting a supernova²) or when a neutron star falls into a black hole³. Reeves *et al.* found that the X-rays emitted in the wake of a GRB come from extremely hot gas expanding outward from the source of the γ -rays and that this gas is highly enriched with the by-products of a supernova explosion. The new results strengthen the link between GRBs and supernovae, and favour models where the GRB occurs within a few days after a supernova.

For some time following their discovery in the late 1960s, bursts of γ -rays were detected by satellites or deep-space probes throughout the Solar System, but their sources were unknown. The directions to most GRBs were very uncertain. Except for a small subclass, called soft γ -ray repeaters, no two bursts with precise locations were ever seen in the same direction in the sky, or identified with any particular class of astronomical object, deepening the mystery of their origins.

During the 1990s, however, a much clearer picture emerged. Today, it is generally accepted that GRBs originate in distant galaxies, rather than in our Galaxy, and are probably the results of the most powerful explosions in the Universe. Most of the alternative models of the birthplace of GRBs were knocked out by a formidable one-two punch. The first blow came from NASA's Burst and Transient Source Experiment on the Compton Gamma Ray Observatory⁴, which showed that GRBs are distributed evenly across the sky rather than being associated with the Milky Way. The second

was in the follow-up observations by the BeppoSAX mission, which indicated that GRBs occur in galaxies at cosmological distances^{5,6}.

The follow-up observations have also found X-ray, optical and sometimes radio 'afterglows' of GRBs that fade over several days. The X-ray afterglows could be evidence of a spherically expanding, cooling fireball following a supernova. But another possibility that has gained favour is that they come from a jet of plasma, moving at relativistic speed — close to that of light — and oriented towards Earth in a conical geometry.

A further puzzle has been that, according to the fireball model, emission lines should be observed, because the elements produced in the supernova are hot enough to excite electronic transitions in the X-ray band, much like sodium atoms in a sodium vapour lamp emit lines in visible (yellow) light. Such lines could provide key diagnostic data about the environments that give rise to GRBs or the physical state of the material that emits γ -rays. These lines have not yet been detected in γ -ray spectra themselves, however. As to the reasons why none have been detected, one can only speculate that the γ -rays result from a nonthermal process after a supernova explosion, or that the emission lines are very weak and the spectral resolution of γ -ray instruments is insufficient to discern them.

Meanwhile, the most powerful X-ray telescopes — NASA's Chandra X-ray Observatory and the European Space Agency's XMM-Newton — have recently obtained evidence of emission lines from iron in the X-ray afterglows^{7,8}. The Doppler widths of the lines (a measure of velocity) give a value of about 10% of the speed of light, giving credence to the expanding-fireball hypothesis.

But these reports have been criticized because the X-ray spectra are very weak and the lines are on the border of detectability. Furthermore, if the lines are identified with iron, there is a mystery as to where the iron originates because it doesn't form in a supernova until after the decay of radioactive ^{56}Ni via ^{56}Co ; and the latter has a half-life of 78 days. So most of the interpretations of the X-ray afterglow spectra have invoked environmental material that was presumably ejected before the explosion. Relativistic jets could also irradiate the ejecta, causing the iron line emission.

Reeves *et al.*¹ have been able to go further in analysing afterglows. Using data gathered by the X-ray spectrometers on the XMM-Newton satellite, they have identified several emission lines in the X-ray spectrum of the afterglow from a GRB that was first observed on 11 December 2001 by BeppoSAX (information about this burst is summarized in ref. 9). The X-ray spectrum of the GRB appears to be different from that of other GRBs and may provide a new context for interpreting the previous X-ray lines. The strongest emission lines Reeves *et al.* detected were from silicon, sulphur and argon; there were other, weaker detections of magnesium and calcium — but no iron was detected. With several different emission lines, especially at lower X-ray energies, one can be much more confident of the identity of these elements.

Reeves *et al.* take full advantage of the additional information to be quarried from these emission lines, including the speed of the gas outflow, its temperature and composition, and the time delay since the supernova explosion. First is the measurement of a Doppler blueshift: the line-emitting gas is moving at about $26,000 \text{ km s}^{-1}$ towards the observer, suggesting that the X-rays are being emitted at the nearest portion of an expanding shell of gas that is illuminated by a narrow cone of γ -rays from the GRB. Next, the lines they detected are highly ionized, indicating that the gas has a temperature of about $5 \times 10^7 \text{ K}$, which is a reasonable value for hot ejecta from a supernova-like explosion.

The data indicate that the material is enriched in these elements, as one might expect in such an explosion where most elements are created and then ejected into the interstellar medium. Reeves *et al.* also put a figure on the size of the expanding shell, estimating it as 10^{15} cm . Given the shell's expansion rate, this gives a time delay from the initial supernova explosion of 10–100 hours, which is consistent with the absence of iron and provides another crucial link to the supernova origin of GRBs.

As with any ground-breaking observation, this one raises questions even as others are answered. For example, why aren't the putative iron emission lines in other GRB

afterglows blueshifted? Why don't lines from lower ionization states appear as the afterglow cools, as would be expected? Also, why don't some of the other GRB afterglows show a similar profusion of elements? The optical and X-ray properties of GRB afterglows can be quite varied^{7,8}. Some are not detected optically, earning them the monikers 'dark GRB' or ' γ -ray bursts hiding an optical source-transient' (GHOST). It may turn out that some of this variation can be explained by the diversity of environments in which the fireballs detonate.

There is evidently a great deal yet to be discovered about GRBs, and two further NASA missions are dedicated to studying them. One (Swift) will be launched next year. The other, the High Energy Transient Explorer, was launched in 2000 and provides accurate GRB positions within minutes to hours so that

other instruments can be trained on the source to catch the afterglows. Continued observations of various types will help to elucidate the origins of GRBs. But it is clear that high-resolution spectroscopy, as exploited by Reeves *et al.*, is fast becoming an especially valuable tool in the endeavour to understand these hugely energetic phenomena. ■

Herman L. Marshall is at the Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

e-mail: hermanm@space.mit.edu

1. Reeves, J. N. *et al.* *Nature* **416**, 512–515 (2002).
2. Meszaros, P. & Rees, M. *Astrophys. J.* **397**, 570–575 (1992).
3. MacFadyen, A. & Woosley, S. E. *Astrophys. J.* **524**, 262–289 (1999).
4. Meegan, C. *et al.* *Nature* **355**, 143–145 (1992).
5. Costa, E. *et al.* *Nature* **387**, 783–785 (1997).
6. Metzger, M. *et al.* *Nature* **387**, 878–880 (1997).
7. Antonelli, L. A. *et al.* *Astrophys. J.* **545**, L39–L42 (2000).
8. Piro, L. *et al.* *Science* **290**, 955–958 (2000).
9. www.aip.de/~jcg/grb011211.html

Stem cells

Cell fusion causes confusion

Andrew E. Wurmser and Fred H. Gage

'Transdifferentiation' is a poorly understood process invoked to explain how tissue-specific adult stem cells can generate cells of other tissues. New results challenge its existence.

Most readers have probably heard of stem cells — unspecialized cells that can generate a variety of more specialized cell types. Not so long ago, it was thought that only embryonic stem cells (ES cells) could generate all the different cell types in the mammalian body¹, and that adult stem cells are more restricted in their developmental potential. Recently, however, many researchers have observed that adult neural and haematopoietic stem cells are not limited to forming cells of the central nervous system or blood, respectively, but can also generate liver, intestinal cells, and heart and skeletal muscle. It has been proposed that they do so by 'transdifferentiating' (acquiring broader developmental potential). If true, this is enormously important — it implies that tissue-specific adult stem cells harbour much, if not all, of the potential of ES cells. This would remove the need to collect stem cells from human embryos for clinical purposes, thus overcoming many of the political and ethical barriers to stem-cell therapy.

Now, however, Ying *et al.*² and Terada *et al.*³ (see pages 542–548 of this issue) raise doubts about whether transdifferentiation actually occurs. The authors set out to show that ES cells can induce neural stem cells² or bone-marrow cell populations containing haematopoietic stem cells³ to transdifferentiate into embryonic-like stem cells when cultured together *in vitro*. Instead, they found that the adult stem cells spontaneous-

ly fused with the ES cells and took on their characteristics — an event that may previously have been misinterpreted as transdifferentiation.

One test of transdifferentiation *in vitro* is based on the idea that the developmental limitations of tissue-specific stem cells are dictated by their environment, and that new signals that relax these restrictions might be provided by cells from a different tissue. In several experiments, for example, neural stem cells were isolated from the central nervous system of mice, labelled with a marker protein such as β -galactosidase or green fluorescent protein (GFP) to make them easy to trace, and cultured with other cell types. Amazingly, when neural stem cells expressing such markers were cultured with myoblasts (precursors of skeletal muscle cells) or embryoid bodies (produced from ES cells), they differentiated into β -galactosidase- or GFP-labelled skeletal myotubes (muscle cells)^{4–6}. Moreover, GFP-labelled neural stem cells differentiated into GFP-labelled heart-muscle cells when mixed with heart-muscle cells from newborn rats⁷. It was concluded that the stem cells achieved this by transdifferentiation^{5,7}, although there are other interpretations (Fig. 1a, overleaf).

The experiments by Ying *et al.*² and Terada *et al.*³ were different in that both starting cell populations were tagged with distinct markers, enabling the authors to trace the fate of each cell type. Ying *et al.*² isolated neural stem cells from mice in which two

genes — one encoding GFP and one encoding a protein that confers resistance to the antibiotic puromycin — are linked to a control region from the *Oct4* gene, which is expressed only in ES cells⁸. These cells were cultured with ES cells that had been genetically engineered to include the hygromycin phosphotransferase gene, which confers resistance to the antibiotic hygromycin. The culture medium was an ES-cell growth medium that contained puromycin, to select against these original ES cells.

After two to four weeks, the authors recovered several cell colonies that expressed GFP and were resistant to puromycin, indicating that they were derived from neural stem cells but had become ES-cell-like with respect to their growth characteristics and the *Oct4*-driven expression of these markers. Taken alone, this might have been interpreted as evidence of transdifferentiation. Surprisingly, however, the cells also expressed hygromycin phosphotransferase (which could only have come from the ES cells); moreover, 18 independently isolated cell lines had twice or nearly twice the normal DNA content. Ying *et al.* conclude that the ES-like cells arose from fusion of the neural and embryonic stem cells, rather than from transdifferentiation (Fig. 1b).

Terada *et al.*³ independently took a similar approach (Fig. 1c) by culturing ES cells with bone-marrow cell populations, rich in haematopoietic stem cells, that expressed GFP and the puromycin-resistance gene. Puromycin was added to eliminate the ES cells, and the signalling molecule interleukin-3 was withdrawn to suppress the growth of bone marrow cells. This resulted in proliferating colonies of GFP-labelled, puromycin-resistant cells that mimicked ES cells in their morphology and growth kinetics, expressed ES-cell marker proteins, and differentiated into heart-muscle cells after the removal of growth factors. These cells, too, might have been thought to arise through transdifferentiation. But Terada *et al.* found that 2 cell lines had 4 sex chromosomes (XXXY, the Y chromosome having come from the ES cells), and 11 had twice the usual amount of DNA, again suggesting that cell fusion had occurred.

Could cell fusion have been misinterpreted as transdifferentiation in previous studies? A complicating factor is that transdifferentiation has been seen to occur in 7–57% of the neural stem-cell population in co-culture assays^{5,6}, whereas Ying *et al.*² and Terada *et al.*³ detected cell fusion at the lower frequency of 1 in 10^4 to 1 in 5×10^5 . The actual rate of fusion may be higher: the stringency of the selection approach and the long culture period required to assay cell fusion (compared with transdifferentiation experiments, which are completed within four to five days^{5,7}) might have resulted in deceptively low levels of detectable cell fusion because

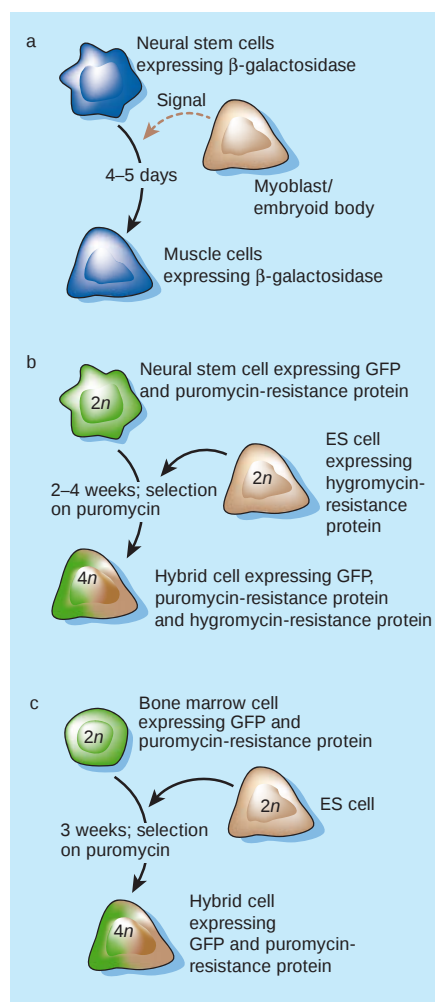


Figure 1 Can adult stem cells transdifferentiate into other cell types? a, In previous work^{4,5} neural stem cells, labelled with β -galactosidase, were cultured with myoblast cells or embryoid bodies, producing β -galactosidase-labelled muscle cells. This was interpreted as evidence that the stem cells received signals that caused them to transdifferentiate into muscle cells. But there are other possibilities. The stem cells might have been contaminated with muscle precursor cells; a few stem cells might have fused with myoblasts or embryoid-body cells. b, Ying *et al.*² cultured puromycin-resistant neural stem cells that expressed green fluorescent protein (GFP) with hygromycin-resistant embryonic stem (ES) cells. After selection, the surviving cells expressed GFP, were resistant to puromycin and hygromycin, and had double the normal amount of DNA (4n versus 2n). This suggests that the cells arose by fusion. c, Terada *et al.*³ cultured GFP-labelled, puromycin-resistant bone marrow cells with ES cells. The resulting cells expressed GFP, were resistant to puromycin and had ES-cell properties. They also had double the usual DNA content.

of cell death. Alternatively, cell fusion may indeed be exceedingly rare, and thus unable to account for transdifferentiation by adult stem cells; moreover, the stringent

approach^{2,3} may have driven cell fusion artificially, as the only means by which cells could survive and proliferate. It remains to be seen whether cells fuse under less selective conditions.

In addition, cell fusion has been detected only *in vitro* so far^{2,3}. But both *in vitro* and *in vivo* observations have been taken as evidence of transdifferentiation. For instance, neural stem cells and bone marrow cells transplanted into mice have been shown to contribute to the central nervous system, blood, liver and muscle (see refs 2, 3). Nevertheless, it remains formally possible that cell fusion or alternative mechanisms of cell conversion may be significant *in vivo*, making it important to re-examine the *in vivo* data. These issues are particularly acute in instances where tissue-specific stem cells convert into multinucleated cells, such as skeletal myotubes and Purkinje neurons^{5,9}.

These questions notwithstanding, it is remarkable that cells can merge at all given how complicated it is to fuse membranes. Intracellular membrane fusion, for example between transport vesicles and the plasma membrane, involves intricate molecular machinery, including proteins known as SNAREs and Rabs¹⁰. The mechanism of cell fusion is unknown, but interactions between plasma-membrane molecules in the fusing cells might tether the cells together, partly mimicking the function of SNARE and Rab proteins.

Finally, an interesting feature of the fused cells^{2,3} was that they had many characteristics of ES cells, suggesting that the embryonic genetic programme may dominate over that of adult stem cells. So cell fusion might afford a means of identifying proteins that make ES cells so versatile. Most important, however, the new results call for a detailed genetic analysis of cells that have been thought to undergo transdifferentiation, and highlight the need to define this process mechanistically so that it can be more rigorously diagnosed in future. Until then, it may be unwise to make generalizations or policies that limit efforts to understand the still-puzzling behaviour of stem cells.

Andrew E. Wurmser and Fred H. Gage are in the Laboratory of Genetics, The Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.
e-mail: gage@salk.edu

- Gardner, R. L. & Beddington, R. S. J. *Cell Sci.* **10** (suppl.), 11–27 (1988).
- Ying, Q.-L., Nichols, J., Evans, E. P. & Smith, A. G. *Nature* **416**, 545–548 (2002); online 13 March 2002 (DOI 10.1038/nature729).
- Terada, N. *et al.* *Nature* **416**, 542–545 (2002); online 13 March 2002 (DOI 10.1038/nature730).
- Clarke, D. L. *et al.* *Science* **288**, 1660–1663 (2000).
- Galli, R. *et al.* *Nature Neurosci.* **3**, 986–991 (2000).
- Rietze, R. L. *et al.* *Nature* **412**, 736–739 (2001).
- Condorelli, G. *et al.* *Proc. Natl Acad. Sci. USA* **98**, 10733–10738 (2001).
- Yeom, Y. I. *et al.* *Development* **122**, 881–894 (1996).
- Priller, J. *et al.* *J. Cell Biol.* **155**, 733–738 (2001).
- Clague, M. J. *Curr. Biol.* **9**, R258–R260 (1999).



100 YEARS AGO

At the request of the U.S. Senate, the Secretary of Agriculture has reported upon the condition of the American bison. In his summary he states that this species is on the verge of extermination. Scarcely a handful now remain of the millions which formerly roamed over the plains of the west. Only two small herds of wild buffalo are in existence in the United States—one in the Yellowstone Park, the other in Lost Park, Colo. There are no wild buffalo in Canada, except in the Peace river country, where a few woodland buffalo, believed to be a different species from the American plains buffalo, still exist. A number of buffalo have been domesticated and half-domesticated, there being three important herds in addition to the small herds in zoological parks and in the hands of private individuals. It is suggested that if the Government would acquire possession of a considerable number of full-blooded animals the absolute extermination of the species might be long delayed.

From *Nature* 3 April 1902.

50 YEARS AGO

The Indian Science Congress has been holding its annual sessions since the year 1914 at different towns and cities of India with the view of enabling eminent men of science of India as well as from abroad to meet on a common platform of intellectual kinship and to keep abreast of the times in discussions of scientific problems. The 1952 session was held early in January in Calcutta... In a spectacular pandal bedecked with floral and artistic designs and packed to its capacity of more than six thousand people, the session was opened by the chanting of Vedic hymns and the recitation of *Vande Mataram*... Then the Honourable Shri Jawaharlal Nehru, Prime Minister of India, delivered an address, in the course of which he called upon scientific workers of every rank and level to lend him their active cooperation in the solution of the gigantic problems that are facing the teeming millions of India. Mentioning that the setting up of national laboratories was a matter of pride for the people of India, he observed that the masses ought to realize the importance of the role that scientific men have to take in the building up of a truly modern India.

From *Nature* 5 April 1952.

Biogeography

Baffled over bison

Peter D. Moore

Two centuries ago, bison and elk were more common to the east of North America's Rocky Mountains than to the west. A look into why this was so may hold lessons for modern conservation biologists.

In 1804, President Thomas Jefferson sanctioned and funded one of the most important expeditions in North America's history. Meriwether Lewis and William Clark set off from St Louis, Missouri, with a commission to survey the upper Missouri River basin and to seek a means of crossing the Rocky Mountains to reach the west coast of the continent (Fig. 1). Their geographical and biological discoveries were of great consequence for the development of the region, but they also brought to light some difficult questions for future biogeographers. One that is still not fully answered is why two large ungulates then common on the east side of the Rockies, namely the bison (*Bison bison*) and the elk (*Cervus canadensis*), were scarce on the western side. Writing in *Conservation Biology*, Lee Lyman and Steve Wolverton¹ take up the 200-year-old challenge posed by Lewis and Clark's observations and conclude that a range of factors, involving both climate and natural and human predation, were involved.

The journals of the expedition, written by both Lewis and Clark, supply data on the contemporaneous distribution and abundance of bison and elk. These include qualitative observations, such as the fact that bison became scarcer as they proceeded west along the upper Missouri River in Montana, and also quantitative data in the form of records of game killed to supply food for the expedition. Their figures relating to game kills have been analysed previously by Martin and Szuter², who came to the conclusion that the lack of game on the western side of the Continental Divide was a consequence of intense hunting by native American peoples, resulting in a 'game sink'. The intensity of such human 'predation' was less on the east of the mountains, they claimed, because this was disputed territory that lay on tribal boundaries, resulting in high game populations within this 'war zone'. But Lyman and Wolverton¹ have re-examined the kill data, and express their results with greater temporal and geographical precision. They supply the numbers of animals killed in each area per day, providing a more accurate picture of the relative abundance of game species over the area covered.

The recalculated data¹ present a rather different picture from that of Martin and Szuter in two respects. First, the boundary between game-rich and game-poor areas is

much more diffuse than was originally proposed; there is no sharp division between game sink and war zone. Second, there is high spatial and temporal variability even within the game-rich region. Some sites that were game-poor on the westward journey proved rich on the return leg. But, as the original journals give no indication of the effort made by the expedition in hunting, even the variability may be spurious. The idea of a clear distinction between rich and poor areas cannot be sustained. Nonetheless, the relative scarcity of game to the west of the Rockies compared with the east is still apparent.

Why should this be? Would elk and bison numbers have been substantially higher in the west if human predation in that region had been lower? This, like all historical hypotheses, is difficult to test. Martin and Szuter² claim that horses and cattle have since managed perfectly well west of the mountains, so grazing resources for herbivores such as elk and bison were evidently available. This is not a very robust argument, however, because large herbivores do not all have the same dietary or ecological requirements³, and Lyman and Wolverton¹ also show that cattle have been successful in the area only when supplied with supplementary food.

Historical theses may be best tested by historical methods. Archaeological data on bison kill sites in the past 10,000 years show a much greater abundance of such sites east of the Rocky Mountains, although there is a scattering to the west. So the mountains were not an insurmountable physical barrier to the bison, and archaeology seems to indicate that there was, in fact, higher human predation in the east than the west.

So factors other than human predation must have limited the success of these large herbivores in the west coast region, and a climatic factor such as high snowfall is one possibility. Daubenmire⁴ has observed that bison can remove light snow (up to 75 cm deep) from buried vegetation with their heads, but that they are defeated by deep snow. As the western region has heavy snowfalls, this could adversely affect bison populations in the area. Under such circumstances, any additional stress, such as human predation, accompanied by limited recolonization through the mountain passes following local extinction, could well account for the bison's lack of success in the



Figure 1 North American explorations. Meriwether Lewis and William Clark's famous expedition brought to light questions that tax biogeographers to this day. (Painting by Alfred Russell.)

west. The Rocky Mountains would thus act as a filter rather than a complete barrier to animal migrations.

Archaeological data also confirm the presence of elk to the west of the Rocky Mountains throughout the past ten thousand years or so, and indicate a sustained level of human predation. But within the last century, elk abundance has greatly increased. Lyman and Wolverton¹ account for this in terms of lower natural predation rather than less human predation. The extirpation of wolves and cougars from much of the region in the early twentieth century is the likely cause of growing elk populations, so natural predation was probably a constraint on population growth in earlier times.

The conclusion of this piece of research into the historical biogeography of the Pacific Northwest is that a combination of

factors, rather than simply human predation, produced the past pattern of abundance of these two ungulates. The historical approaches used here also have much to offer conservation biologists. If the factors that have limited biogeographical distributions and population levels of organisms in the past are better understood, then current programmes of reintroduction and management can be more effectively evaluated. The past, for the modern ecologist, often provides a key to the present.

Peter D. Moore is in the Division of Life Sciences, King's College, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NN, UK. e-mail: peter.moore@kcl.ac.uk

1. Lyman, R. L. & Wolverton, S. *Conserv. Biol.* **16**, 73–85 (2002).
2. Martin, P. S. & Szuter, C. R. *Conserv. Biol.* **13**, 36–45 (1999).
3. Menard, C., Duncan, P., Fleurance, G., George, J.-Y. & Lilas, M. *J. Appl. Ecol.* **39**, 120–133 (2002).
4. Daubenmire, R. *Ecology* **66**, 622–624 (1985).

Telecommunications

A boost for fibre optics

Z. Vally Vardeny

A possible solution to the need to expand fibre-optic capacities presents itself in the form of a new way of amplifying the information-carrying light transmissions. But turning principle into practice will require more work.

In fibre-optic communications, a modulated light beam is sent through a fibre made of silica glass. The modulation carries the information to be transmitted, and the light beam is confined inside the fibre core by the large difference in refractive index between the core and a cladding layer. A fundamental

limiting factor is attenuation or absorption of the information-carrying light beam^{1,2}. In consequence, fibre optics are restricted to use of infrared wavelengths between 1.3 and 1.5 μm , because in this spectral range the losses due to light absorption and scattering, as well as dispersion in the silica fibre, are minimal.

There is, however, an insatiable demand for communication channels, mainly driven by the Internet industry. The infrared range will soon be saturated, so other wavelengths will have to be considered^{3,4}. To overcome attenuation, the light will have to be boosted at points along the fibre, and this could be achieved with laser amplifiers. Writing in *Applied Physics Letters*, Dou *et al.*⁵ describe just such a device — a polymer micro-ring laser — that could, in principle, provide amplification in the visible spectral range when combined with new types of fibre-optic cables.

A laser consists of a 'gain' medium, here a polymer, in which 'population inversion' of the energy levels has to take place to induce stimulated emission of a coherent light beam. Stimulated emission is achieved by exciting the medium by injecting either photons (optical pumping) or electrons (electrical pumping). The laser action, or more precisely the stimulated emission, boosts the intensity of the information-bearing beam.

The optical-gain medium used by Dou *et al.*⁵ is a light-emitting polymer (a dye-doped poly(methylmethacrylate), PMMA), which was coated in a ring onto a glass-fibre core 63 μm in diameter. Other light-emitting coatings — for example π -conjugated polymers such as poly(*p*-phenylene-vinylene), PPV, which strongly emits light without the need for dye-molecule dopants — have previously been used in the same way^{6–8}. What is different about the work of Dou *et al.* is that they have excited the polymer ring from the inside, with an excitation light beam travelling along the optical fibre. This technique is known as longitudinal pumping (Fig. 1a, overleaf). In the earlier work^{6–8} with polymer micro-ring lasers, the gain medium was optically excited from the outside, by transverse pumping (Fig. 1b).

Dou *et al.* claim that longitudinal optical excitation occurs because peripheral waves of the excitation beam reach and stimulate the polymer gain medium. They compare transverse and longitudinal optical excitations of the same gain medium and show that the lasing threshold for longitudinal excitation is lower by a factor of about 40. This has a tremendous advantage: high power is required for transverse excitation, and the resulting heat can distort the polymer ring around the fibre.

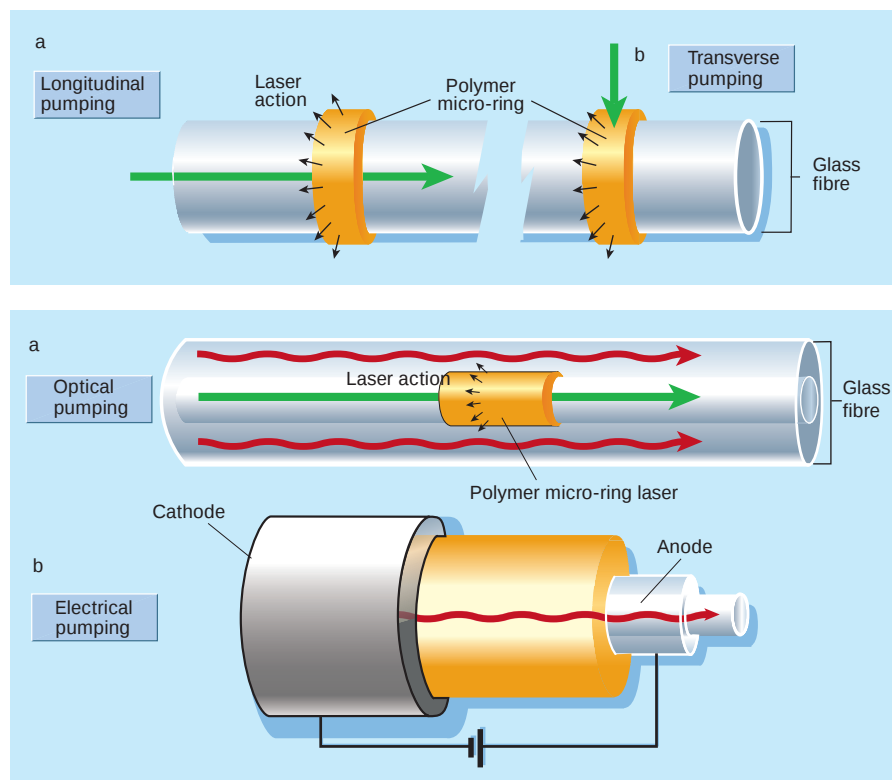
So far so good: we have the highly attractive prospect of using longitudinal stimulation of polymer micro-ring lasers for fibre-optic amplification. But new designs of fibre will be required to put the principles outlined by Dou *et al.*⁵ into practice. The authors themselves experimented with optical pumping of the gain medium, but the same end might also be achieved by electrical means.

Figure 2 shows possible fibre designs for

Figure 1 (top) **Pumping polymer micro-ring lasers.** Pumping is necessary for laser action and ensuing amplification of the modulated (information-carrying) light beam in fibre-optic communications using light at visible wavelengths between 0.4 and 0.7 μm . The modulated beam is not shown here.

a, Longitudinal pumping, which is generated by a beam (green) travelling along the optical fibre. b, Transverse pumping, which is produced with an external light source. Dou *et al.*⁵ show that longitudinal pumping has considerable advantages.

Figure 2 (bottom) **Speculative schemes for longitudinal pumping of polymer micro-ring lasers.** a, Optical pumping. Here the excitation beam (green) is confined in an inner, single-mode fibre that is surrounded by the polymer micro-ring amplifier. The information-carrying beam requiring amplification (red) travels along the multimode cladding layer. b, Electrical pumping. The information-carrying beam (red) is amplified by a surrounding polymer medium, which is pumped by electrical flow between two electrodes: a semi-transparent anode (made of 'TTO') and an aluminium cathode.



optical and electrical pumping. In Fig. 2a (optical pumping), two light beams are transmitted. The inner, excitation beam optically stimulates a surrounding polymer micro-ring laser, which then amplifies the modulated information-bearing beam. As depicted here, the optical-fibre system might consist of a single-mode fibre for the excitation beam, surrounded by a multimode cladding layer for the modulated beam. Or the system could instead use the central hole in a newly invented type of photonic crystal fibre^{3,4}, in which the polymer layer would be deposited on the inside surface of the hole.

The alternative — electrical stimulation of the optical-gain medium — is depicted in Fig. 2b. In this pumping scheme, the polymer is sandwiched between two cylindrical electrodes, and positive and negative charges are injected into the polymer active layer. In practice, electrically stimulated amplification rings would have to be stationed at intervals along the fibre. Other studies⁹ with an organic diode, using PPV as the active layer, have already shown that light emission with cylindrical geometry is possible, and laser action in the form of stimulated emission should also be feasible. Alternatively, other organic compounds with emission in the infrared could be used in cylindrical configurations to boost light at 1.3–1.5- μm wavelengths, both optically¹⁰ and electrically¹¹, in present fibre-optic systems.

For the moment, these amplification schemes remain pie in the sky. Nonetheless, the principles explored by Dou *et al.*⁵ are promising — and one way or the other we

will have to find ways to expand the capacities of fibre-optic systems.

Z. Vally Vardeny is in the Department of Physics, University of Utah, Salt Lake City, Utah 84112, USA.

e-mail: val@physics.utah.edu

1. Glass, A. M. *et al.* *Bell Labs Tech. J.* 168–187 (January 2000).
2. Ainslie, J. & Day, C. R. *J. Lightwave Technol.* 8, 967–979 (1986).

3. Birks, T. A. *et al.* *Opt. Lett.* 22, 961–963 (1997).
4. Cregan, R. F. *et al.* *Science* 285, 1537–1539 (1999).
5. Dou, S. X. *et al.* *Appl. Phys. Lett.* 80, 165–167 (2002).
6. Kuwata-Gonokami, M. *et al.* *Opt. Lett.* 20, 2093–2095 (1995).
7. Frolov, S. V., Shkunov, M. & Vardeny, Z. V. *Phys. Rev. B* 56, R4363–R4366 (1997).
8. Kawabe, Y. *et al.* *Appl. Phys. Lett.* 72, 141–143 (1998).
9. Fujii, A., Frolov, S. V., Vardeny, Z. V. & Yoshino, K. *Jpn. J. Appl. Phys.* 37, L740–L742 (1998).
10. Gillin, W. P. & Curry, R. J. *Appl. Phys. Lett.* 74, 798–799 (1999).
11. Tessier, N. *et al.* *Science* 295, 1506–1508 (2002).

Genetics

Immaculate misconception

M. A. Surani

In mammals, mother and father make an equal genetic, but an unequal 'epigenetic', contribution to offspring. Studies of humans and mice with no maternal epigenetic contribution reveal more about this asymmetry.

William Harvey, an anatomist and personal physician to two kings of England, was taking a gamble when he proposed in 1651 that 'Ex ovo omnia' — 'everything comes from an egg'. It wasn't until much later that the mammalian egg, or oocyte, was first detected and Harvey was proved right: a whole organism can develop from this remarkable cell. But what he did not know is that an input from sperm is also essential for mammalian oocytes to fulfil this potential. This is, of course, why sex is necessary and 'virgin' births are impossible in humans. The need for sperm lies in the fact that all genes are not equal. A fertilized egg contains two copies of every gene, one from sperm and one from oocyte. Usually, both

copies are expressed in relevant cells in the embryo. But some genes are specifically labelled ('imprinted') so that only the maternal or paternal copy is active. So an embryo receives an equal genetic contribution from its parents, but the parental genomes are not functionally equivalent, explaining why both are needed. Indeed, embryos with two male or two female genomes cannot develop normally.

Three new papers^{1–3} provide another graphic demonstration of the importance of these unequal 'epigenetic' contributions from sperm and oocyte. On page 539 of this issue¹, Judson and colleagues describe a unique case in which a mutation seems to prevent human oocytes from acquiring any

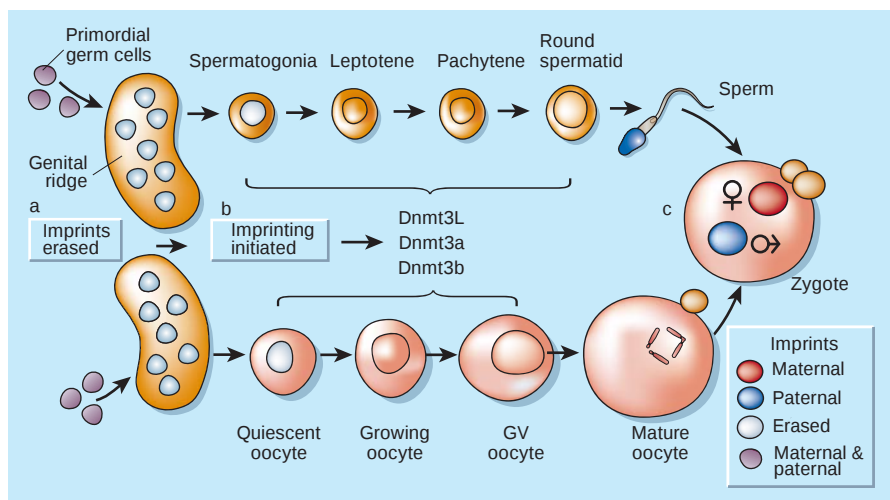


Figure 1 **Developing imprinting.** Imprinted genes are those that are 'marked' in a parental-specific way. **a**, As primordial germ cells enter the genital ridge of developing female and male embryos, all imprinting marks are erased¹⁰. **b**, Bottom, as female germ cells grow and develop into oocytes, the *Dnmt3L* gene, possibly with *Dnmt3a* and/or *Dnmt3b*, is involved in reintroducing maternal-specific imprints^{2,3}. Top, these genes are also expressed during sperm development but how they function in paternal imprinting is unknown. **c**, After fertilization, a zygote is formed with both genetic and epigenetic (imprinting) contributions from the parental genomes. GV oocyte, germinal-vesicle oocyte.

maternal-specific imprints at all. The mutant eggs fail to develop even after fertilization. Meanwhile, Bourc'his *et al.*² (writing late last year in *Science*) and Hata *et al.*³ (in a report in *Development*) have described a remarkably similar outcome in mice with a mutation in the *Dnmt3L* gene, which is required for imprinting in mouse oocytes.

Nearly 50 mammalian imprinted genes have been identified so far, and their significance for human development is clear from Judson *et al.*'s study¹ of a woman in whom conceptuses (very early embryos) developing from normally fertilized eggs invariably died soon after they implanted into the wall of the uterus. Judson *et al.* analysed one such conceptus, and found that it had a normal genetic contribution from sperm and oocyte. The problem, it turns out, was with the epigenetic status of the oocyte: all the maternally inherited epigenetic marks (recognized as methyl groups on cytosine-guanine base pairs in particular DNA sequences associated with imprinted genes) were missing from the conceptus. Paternally inherited epigenetic marks were normal.

The physical characteristics of the con-

ceptus were also revealing — they were similar to those that occasionally arise spontaneously after the embryonic loss of the maternal genome, leaving only the paternal genome. Why should a conceptus that lacks maternal imprints develop as if it lacks a maternal genome entirely? The reason is that, during the development of primordial germ cells, from which oocytes and sperm will form, all imprints are first erased (Fig. 1). New parental-specific imprints are introduced subsequently when sperm and oocytes begin to mature⁴, with oocytes receiving most known imprints⁵. So a maternal genome that lacks maternal imprints functions more like a paternal genome in the fertilized egg — although not entirely^{4,6}. Although relatively sparse, the epigenetic marks in sperm are essential, as a conceptus with two maternal genomes also fails to develop.

How, then, are imprints laid down in developing oocytes? Here the studies by Bourc'his *et al.*² and Hata *et al.*³ are important, as they show that the *Dnmt3L* gene is crucial to this process in mice. Bourc'his *et al.*² engineered animals with a targeted mutation in both copies of *Dnmt3L*. The mutation

had no immediate effect on development: the original, 'founder' males and females matured to adulthood, although the males did not produce sperm and were sterile.

The females lacking functional *Dnmt3L* protein did produce oocytes. But, after fertilization of these oocytes by normal sperm, conceptuses failed to develop after implantation, and showed several embryonic and placental abnormalities. These defects were clearly of maternal origin, as the embryos lacked all maternally inherited epigenetic modifications but, of course, not those from sperm². Some of the abnormalities, for example in the placenta, were similar to those seen in conceptuses with paternal genomes only, or with genomes that lack all parental imprints^{4,6}. Furthermore, because of the lack of maternal imprints, several imprinted genes were regulated abnormally, with some being repressed; this probably contributed to the observed abnormalities².

So it seems that *Dnmt3L* is crucial in laying down imprints in the female germ line, at least in mice. But in the human conceptus studied by Judson *et al.*¹, there was no identifiable mutation in the human *DNMT3L* gene (D. Bonthron, personal communication). Mutations in another gene must account for their observations.

Two possibilities are the *DNMT3A* and *DNMT3B* genes, which share similarities with *DNMT3L*. The encoded proteins are two DNA methyltransferase enzymes, which can initiate *de novo* DNA methylation^{2,7,8}. Indeed, they might be required for *DNMT3L* to work, as *DNMT3L* does not encode the catalytic domain required for DNA methylation. Consistent with this idea, Hata *et al.*³ show that the mouse *Dnmt3L* protein can interact with these enzymes, and that the absence of *Dnmt3a*, like the lack of *Dnmt3L*, results in the loss of maternal imprints in mice. But the methyltransferases cannot themselves detect specific DNA sequences that require *de novo* methylation, so some other modification might be needed first. This might involve the methylation of histone proteins — part of the packaging that enables DNA to be squeezed into the nucleus. For example, an interaction between methylated histone H3 protein and another packaging protein, HP1 β , can lead to *de novo* DNA methylation⁹.

Whatever the answers, the new papers¹⁻³ stress the importance of the epigenetic asymmetry between parental genomes for normal mammalian development¹⁰ (Fig. 2). Even a slight deviation in the epigenetic marks on some chromosomal regions, such as on chromosomes 11 or 15, can result in devastating human diseases such as Beckwith-Wiedemann and Prader-Willi/Angelman syndromes¹¹. Moreover, the cloning of mice and other mammals is possible because the donor nuclei, derived from adult tissues, possess the appropriate parental imprints¹⁰.

Why imprinting exists at all remains

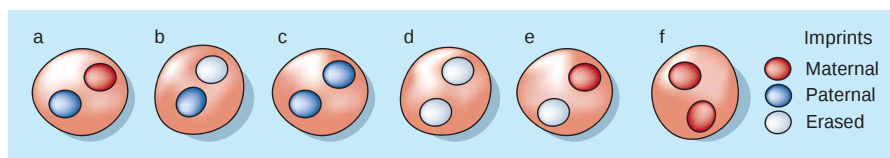


Figure 2 **The importance of imprinting.** **a**, Epigenetic asymmetry between paternally and maternally derived genomes, as shown here, is essential for normal development. **b**, The absence of the *Dnmt3L* and probably *Dnmt3a* genes in developing mouse oocytes prevents maternal-specific imprinting^{2,3}. After fertilization, the zygote contains a maternal genome with no imprints and a normal paternal genome; it cannot develop normally. Judson *et al.*¹ have described a similar outcome in a human patient, although the genetic basis is unclear. **c-f**, These combinations of genomes in different epigenetic states are also incompatible with full development⁴.

enigmatic. There are signs that, during evolution, a war broke out between the sexes that left the maternal genome with most of the DNA-methylation marks associated with imprinted genes⁵. Even now we observe the preferential demethylation of the paternal genome after fertilization; the maternal genome remains largely unaffected¹². This suggests that maternally inherited proteins in the oocyte are used to demethylate, and so regulate the function of, the paternal genome after fertilization. Nevertheless, even the relatively few epigenetic marks that have survived in the paternal genome are crucial for development and hence a significant barrier to virgin birth. So for the time being men are indispensable. But with the increasing pace of research, even this barrier might be breached in the future. I for one will

be following the unfolding story with fascination — and apprehension. ■

M. A. Surani is at the Wellcome/Cancer Research UK Institute of Cancer and Developmental Biology, Tennis Court Road, Cambridge CB2 1QR, UK. e-mail: as10021@mole.bio.cam.ac.uk

1. Judson, H., Hayward, B. E., Sheridan, E. & Bonthron, D. T. *Nature* **416**, 539–542 (2002).
2. Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B. & Bestor, T. H. *Science* **294**, 2536–2539 (2001).
3. Hata, H., Okano, M., Lei, H. & Li, E. *Development* (in the press).
4. Obata, Y. *et al. Development* **125**, 1553–1560 (1998).
5. Reik, W. & Walter, J. *Nature Genet.* **27**, 255–256 (2001).
6. Kato, Y. *et al. Development* **126**, 1823–1832 (1999).
7. Aapola, U. *et al. Genomics* **65**, 293–298 (2000).
8. Okano, M., Bell, D. W., Haber, D. A. & Li, E. *Cell* **99**, 247–257 (1999).
9. Tamaru, H. & Selker, E. U. *Nature* **414**, 277–283 (2001).
10. Surani, M. A. *Nature* **414**, 122–128 (2001).
11. Nicholls, R. D. *et al. J. Clin. Invest.* **105**, 423–418 (2000).
12. Mayer, W. *et al. Nature* **403**, 501–502 (2000).

Developmental biology

The plastic face

Massimo Pasqualetti and Filippo M. Rijli

In vertebrates, the face is formed in part by neural crest cells. It has been assumed that the developmental fate of these cells is inbuilt. New work, however, reveals a role for instructive signals from nearby cells.

Like many developmental processes, the formation of the face in vertebrates depends on the migration of cells from one part of the embryo to another. The main source of cells that contribute to the facial structures is the neural crest¹ — a ridge of embryonic tissue generated by the developing brain. Different bones in the head are formed from different populations of neural crest cells, according to the cells' position along the head-to-tail (anterior–posterior) axis of the embryo. So, for example, anterior neural crest cells migrate to form the brain-case and the nose and jaw bones. Neural crest cells from the posterior part of the brain help to make the neck bones. But beyond these general details, little is known about how facial structures are induced to develop at the right place and in the appropriate, species-specific, shape. Such developmental instructions might be intrinsic to the neural crest and fixed before migration. Alternatively, neural crest cells might be instructed by surrounding tissues during or after migration.

An early, landmark study of chick embryos² favoured the first of these models. But two new papers, published by Trainor *et al.*³ and Couly *et al.*⁴ in *Science* and *Development*, respectively, come down firmly in favour of the second. The two groups find that the neural crest cells that contribute to the face are developmentally 'plastic' — they do not themselves carry any information about their fate, but must be instructed by signals from other tissues to generate skeletal elements of appropriate shape and polarity.

The authors also identify two sources of these signals.

During early vertebrate development, the neural crest forms from the neural tube — a long structure that will generate the brain and spinal cord, and which runs along the back of the embryo (Fig. 1a, overleaf). The part of the neural tube that gives rise to the brain is divided into segments — the forebrain, midbrain and hindbrain; the hindbrain is in turn subdivided into smaller segments called rhombomeres. Neural crest cells generated from the forebrain, midbrain and the first two rhombomeres migrate to cover the brain, as well as into the forehead region and first pharyngeal arch, and contribute to the brain-case, nose and jaws. By contrast, cells from the remaining rhombomeres fill the second, third and fourth arches, which form the neck. (Pharyngeal arches are the embryonic structures from which most cranial features develop.)

Nearly two decades ago, Noden² made a seminal discovery while investigating facial development. Using chick embryos, he replaced neural crest cells that were destined for the second pharyngeal arch with quail cells destined for the first arch. This resulted in the formation of extraneous jaw- and beak-like structures in the neck region — in other words, the grafted cells migrated to the nearest (second) arch, from which the neck forms, but gave rise to the morphology characteristic of the first arch. At face value, this implies that the fate of neural crest cells is set while they are still in the neural tube, deter-

mining which structures they help to form. But Noden pinpointed two potential contradictions. First, the grafted neural crest cells also contributed to normal second-arch-derived structures; and second, the extra skeletal elements invariably included jaw features but never forehead bones.

How can these results be reconciled? Trainor *et al.*³ provide a solution. These authors carried out similar grafting experiments, and found that extraneous jaw structures were produced only when the isthmus (the boundary between the midbrain and hindbrain) was included with the grafted cells. The isthmus is a source of signalling molecules, particularly fibroblast growth factor-8 (FGF8), and normally represses the *Hoxa2* gene — the main determinant of the fate of second-arch cells^{5–8} — in the first rhombomere⁹. Trainor *et al.* found that when the isthmus was transplanted posteriorly into subsequent rhombomeres, it inhibited *Hoxa2* in adjacent second-arch cells, which normally contribute to the neck but now develop as extra jaw bones. Adding FGF8 alone also inhibited *Hoxa2* in neural crest cells. But when Trainor *et al.* moved just the neural crest cells that were destined for the first arch, without the FGF8-secreting isthmus, the cells took on a second-arch fate. In other words, they were not irreversibly committed to their fate before migrating to the arches.

So it seems that signals from the neural tube inform neural crest cells about their original position. But where does information about the shape and polarity of facial bones come from? Couly *et al.*⁴ provide evidence in favour of a tissue — the endoderm — that lies beneath the neural crest and lines the pharyngeal arches.

Using chick embryos again, these authors first excised all the neural crest cells that contribute to the face. They then showed that any short fragment of the region of the neural crest that does not express *Hox* genes can generate the entire facial skeleton. The findings again show that these cells are not endowed with intrinsic developmental information before migration. Couly *et al.* further discovered that distinct regions of the endoderm (regions I–IV; Fig. 1a) instruct *Hox*-negative neural crest cells to produce specific face and jaw bones. The removal of single regions of the endoderm prevented the formation of distinct subsets of facial neural-crest-derived elements. However, this might simply reflect the removal of a general signal needed for skeleton formation. More compellingly, individual stripes of endoderm grafted near their normal counterparts led to spectacular duplications of facial cartilage and adjacent bones (Fig. 1b), with orientations that depended on the orientation of the graft.

These results^{3,4} suggest that distinct populations of neural crest cells respond to

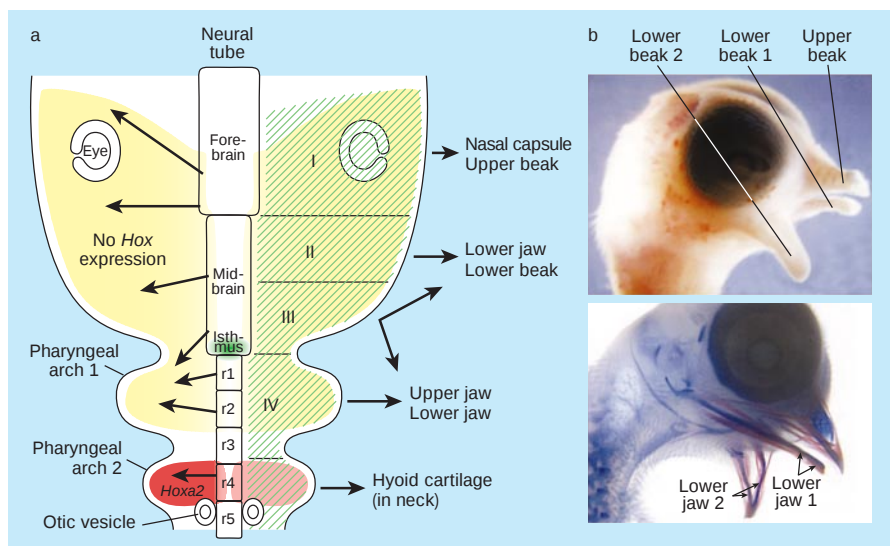


Figure 1 Shaping the face. **a**, Flattened view of the back of a chick embryo. Left, during early development, neural crest cells migrate out of the neural tube and towards the forehead and nasal region (topmost bulge) and the pharyngeal arches. For simplicity, migration is shown only on the left, although it would occur on both sides in the embryo. Right, spatial relationships between the neural crest and the underlying pharyngeal endoderm, which is divided into regions I–IV. Trainor *et al.*³ show that signals from the neural tube (such as from the isthmus) inform neural crest cells destined for the first pharyngeal arch about their position (for example, through repressing the *Hoxa2* gene). However, the cells' fate is not irreversibly set before migration. Then, signals from the underlying endoderm (green stripes) instruct the cells about the size, shape and orientation of the skeletal elements they will form, as Couly *et al.*⁴ show. **b**, Grafting of an extra endodermal region II onto the original yields a fully duplicated lower beak and jaw, highlighting the importance of signals from the endoderm. Modified with permission from ref. 4. **r**, rhombomere.

regional cues from the neural tube (such as FGF8 from the isthmus). Such signals must also be maintained in the arch, possibly to ensure that the cells segregate and migrate correctly. But the cells remain uncommitted until further instructed by the pharyngeal endoderm, each region of which contains information that guides the size, shape and orientation of parts of the facial skeleton. It will be important to identify the factors that control regionalization of the endoderm, and the exact signals that instruct the neural crest. During vertebrate evolution, the distribution or timing of such factors or signals may have been changed so as to mould different facial morphologies.

Some past results may need revisiting in light of the new work. For example, grafting rhombomeres 1 and 2 plus the isthmus more posteriorly in chick embryos^{2,3} — leading to the inhibition of *Hoxa2* — and knocking out the *Hoxa2* gene in mice^{5,8} resulted in jaw structures in the neck region. Given Couly *et al.*'s discovery⁴ that the pharyngeal endoderm instructs neural crest development, this might mean that the endoderm of the second arch provides a signal like that from the first arch. But in that case it seems odd that posterior grafts of only the dorsal parts of rhombomeres 1 and 2, or of rhombomeres 1 and 2 without the isthmus, result in a largely normal second-arch morphology^{3,10} even in the continued absence of *Hoxa2* in the neural crest cells¹⁰.

A related puzzle is why *Hoxa2*-expressing neural crest cells, induced to migrate into the first arch, are unable to generate any skeletal elements at all^{6,10}. A clue might come from an experiment in which jaw-to-neck transformations could be produced only when *Hoxa2* was artificially expressed in both the neural crest and the corresponding (first) arch^{6,7}. Perhaps, to form the neck bones, *Hoxa2* is required not only in neural crest cells but also in other arch components. Tissue-specific knockouts of *Hoxa2* in mice, and further grafting experiments in chicks, may lead to the answers.

Massimo Pasqualetti and Filippo M. Rijli are at the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/ULP, Collège de France, BP 163-67404 Illkirch Cedex, CU de Strasbourg, France.
e-mail: rijli@igbmc.u-strasbg.fr

1. Le Douarin, N. M. *The Neural Crest* (Cambridge Univ. Press, 1983).
2. Noden, D. M. *Dev. Biol.* **96**, 144–165 (1983).
3. Trainor, P. A., Ariza-McNaughton, L. & Krumlauf, R. *Science* **295**, 1288–1291 (2002).
4. Couly, G., Creuzet, S., Bennaceur, S., Vincent, C. & Le Douarin, N. M. *Development* **129**, 1061–1073 (2002).
5. Gendron-Maguire, M., Mallo, M., Zhang, M. & Gridley, T. *Cell* **75**, 1317–1331 (1993).
6. Grammatopoulos, G. A., Bell, E., Toole, L., Lumsden, A. & Tucker, A. S. *Development* **127**, 5355–5365 (2000).
7. Pasqualetti, M., Ori, M., Nardi, I. & Rijli, F. M. *Development* **127**, 5367–5378 (2000).
8. Rijli, F. M. *et al.* *Cell* **75**, 1333–1349 (1993).
9. Irving, C. & Mason, I. *Development* **127**, 177–186 (2000).
10. Couly, G., Grapin-Botton, A., Coltey, P., Ruhin, B. & Le Douarin, N. M. *Development* **125**, 3445–3459 (1998).

Daedalus

Wave to the Sun

Photoelectric solar panels are stiff and brittle. Typically, they are optimized for high efficiency in spacecraft, where they unfold as rigid panels. Even earthbound photoelectric cells share this brittleness. Nature's photo-energetic devices — leaves — are flexible, and much better adapted to the atmospheric life. So DREADCO engineers are inventing a flexible photoelectric cell. Daedalus recalls Ovshinsky glass made to become semiconducting like silicon, the strong boron filaments created by decomposing boron hydride on thin hot wires, and the glass cloth woven from narrow flexible fibres. A narrow wire might be given a fine silicon coating by heating it in a silicon hydride gas. A metallized outer return lead would complete a filamentary photodiode, and the whole thing could be woven into some sort of textile.

DREADCO's 'Photofabric' would probably be inefficient. But as bunting, flags, sails, balloon fabric and so on, it would be well adapted to the atmosphere. Domestic users would love to 'fly the flag' at an energetic profit. Photofabric might even generate bulk electricity. Solar energy meets its main problems a few kilometres up, where clouds absorb sunlight. So Daedalus is inventing his 'kalloon', a combined kite and balloon, made of Photofabric and designed to get above this blocking layer.

Cloud-tops are very sharp when seen from the air. This suggests to Daedalus that the density of the atmosphere must fall at that height. His kalloon will nestle in this discontinuity. One old scheme proposes satellites carrying solar cells, to beam microwaves down through the cloud cover. Daedalus's kalloon will also deliver microwaves back to Earth. Cunningly, he will use its control lines as a spaced double Lecher line. Much lighter than any electrical conductor, this could steer energy very efficiently. A computer at the base of the system, like a tireless kite flier, would reel the lines in or out to keep the kalloon floating on the clouds and facing the moving Sun.

The kalloon might claim some government development money; it captures renewable resources. But Photofabric itself could benefit us all as a small-scale, universal power source. It would provide some of our own power, and thus reduce our power bills, for some folk perhaps to zero. It might change the whole assumption that power, consumed in small quantities, must be generated in big central stations.

David Jones

Organized assembly of carbon nanotubes

Cunning refinements help to customize the architecture of nanotube structures.

Nanoscale structures need to be arranged into well-defined configurations in order to build integrated systems. Here we use a chemical-vapour deposition method with gas-phase catalyst delivery to direct the assembly of carbon nanotubes in a variety of predetermined orientations onto silicon/silica substrates, building them into one-, two- and three-dimensional arrangements. The preference of nanotubes to grow selectively on and normal to silica surfaces forces them to inherit the lithographically machined template topography of their substrates, allowing the sites of nucleation and the direction of growth to be controlled.

Although carbon nanotubes promise to have a wide range of applications¹, better control is needed over the building and organization of nanotube-based architectures. Chemical-vapour deposition (CVD) has been used to align nanotubes vertically on catalyst-printed substrates^{2–8}, and we are now able to control the growth of aligned nanotubes in several directions at once in a single process.

We designed and built all of the nanotube structures on patterns created on silica (SiO₂) and silicon surfaces. The substrates consisted of Si(100) wafers capped

with 100-nm-thick thermal oxide, or thick chemical-vapour-deposited silica layers (up to about 8.5 μm thick). Patterning of Si/SiO₂ was generated by photolithography followed by a combination of wet and/or dry etching.

The catalyst material, which is generally required for CVD growth of nanotubes, was not patterned on the substrates; preparation of the template was therefore simplified. Instead, we stimulated CVD growth of nanotubes by exposing the substrate to a xylene/ferrocene (C₈H₁₀/Fe(C₅H₅)₂) vapour mixture at around 800 °C. This precursor combination gives rise to selective growth of multiwalled nanotubes of diameter 20–30 nm on silica surfaces⁹. There is no nanotube growth on silicon, but the aligned nanotubes grow readily on SiO₂ in a direction that is normal to the substrate surface. The template selectivity for growth is retained down to micrometre-sized SiO₂ templates, allowing ordered nanotube structures to be assembled on Si/SiO₂ substrates.

Figure 1 shows some striking examples of organized nanotube patterns grown on preselected substrate sites. In Fig. 1a, three blocks of micrometre-sized, cylindrical pillars of perfectly vertical, densely packed

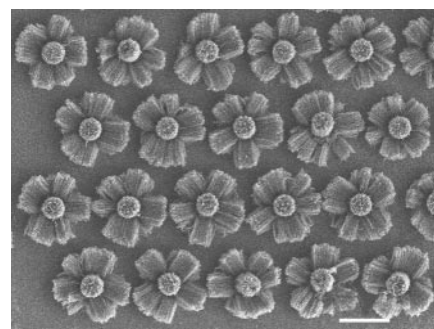


Figure 2 Repeating patterns containing mutually orthogonal nanotube arrays produced on deep (about 5 μm) silica features (circular cross-section) machined on silicon substrates. Growth in the vertical direction occurs from the top silica surface (seen as arrays emanating from the centre of each pattern); growth on the sides occurs as horizontal arrays (sideways growth seen on each pattern). Scale bar, 50 μm.

nanotubes are seen growing from the underlying SiO₂ pattern. The packing of the pillars in each block is determined by the separation between the SiO₂ patterns, and their height is controlled to within about 1–2 μm by tuning the CVD time (growth rates are typically around 10 μm min^{−1}). Nanotube structures with different architecture, such as blocks of rectangular platelets, can be constructed by varying the design of the SiO₂ substrate pattern.

This technique can be used to grow nanotubes in several different directions at once, which was not previously possible. We have realized simultaneous, multidirectional nanotube growth on templates with deeply etched trenches (drilled down to the silicon substrate) separating micrometre-tall SiO₂ features. Figure 1b shows periodic arrays of vertically and horizontally aligned nanotubes grown on repeating patterns of SiO₂ features machined on a planar silicon wafer; the nanotubes are evident in orthogonal directions in the plane of the substrate itself. Beautiful patterns of multiply orientated, organized nanotube structures can be created in this way (Fig. 2).

The simultaneous integration of ordered, geometrically varied nanotube structures in different orientations onto a single substrate could be useful in the manufacture of electromechanical devices. Our fabrication method can be scaled up to large areas and is compatible with standard silicon microfabrication technology.

B. Q. Wei*, R. Vajtai*, Y. Jung*, J. Ward*, R. Zhang†, G. Ramanath*, P. M. Ajayan*

*Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA

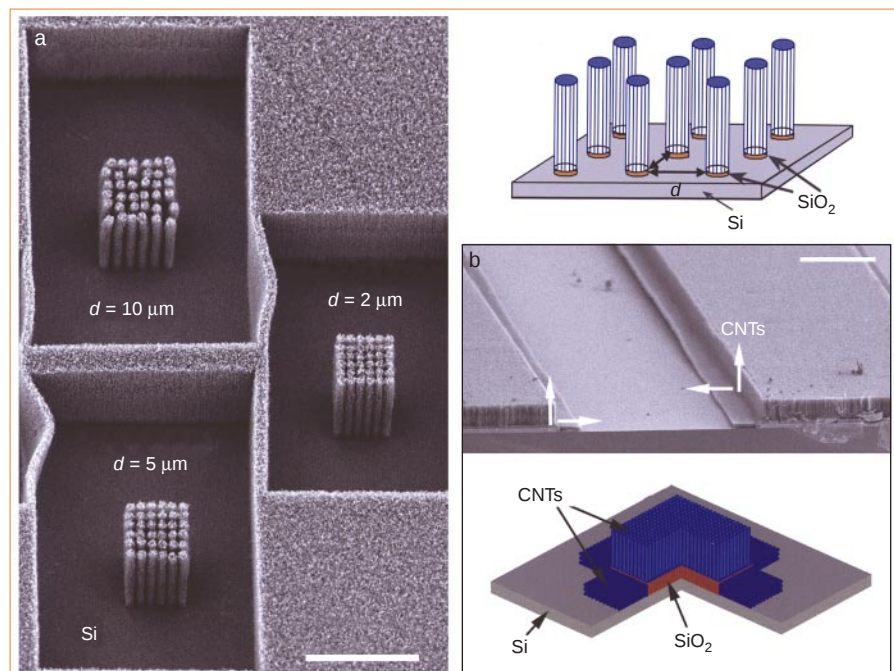


Figure 1 Directed assembly of organized, multiwalled carbon-nanotube structures grown by chemical-vapour deposition. **a**, Image obtained by scanning electron microscopy of three blocks of cylindrical pillars (about 10 μm in diameter) of aligned carbon-nanotube arrays. Each pillar consists of several tens of nanotubes grown in vertical alignment and in a normal direction to SiO₂ patterns on the Si/SiO₂ template. No growth occurs on the Si parts of the template. The separation (*d* in diagram, top right) between pillars in the three blocks is indicated. **b**, Vertical and horizontal growth of aligned nanotubes (CNTs), viewed in a cross-section of a patterned Si/SiO₂ wafer. Scale bars, 100 μm.

e-mail: ajayan@rpi.edu

†Motorola Physical Science Research Laboratories,
Tempe, Arizona 85284, USA

1. Dresselhaus, M. S., Dresselhaus, G. & Avouris, P. *Carbon Nanotubes: Synthesis, Structure, Properties, and Applications* (Springer, Heidelberg, 2001).
2. Li, W. Z. *et al. Science* **274**, 1701–1703 (1996).

3. Sen, R., Govindaraj, A. & Rao, C. N. R. *Chem. Phys. Lett.* **267**, 276–280 (1997).
4. Terrones, M. *et al. Nature* **388**, 52–55 (1997).
5. Ren, Z. F. *et al. Science* **282**, 1105–1107 (1998).
6. Fan, S. S. *et al. Science* **283**, 512–514 (1999).
7. Kind, H. *et al. Adv. Mater.* **11**, 1285–1289 (1999).
8. Schlittler, R. R. *et al. Science* **292**, 1136–1139 (2001).
9. Zhang, Z. J., Wei, B. Q., Ramanath, G. & Ajayan, P. M. *Appl. Phys. Lett.* **77**, 3764–3766 (2000).

Sperm competition

Motility and the midpiece in primates

In animals with multiple-partner mating systems, the gametes of two or more males must compete to fertilize a given set of ova^{1–5}. Here we show that the volume of the midpiece in individual sperm is significantly greater in primate species in which the females mate with multiple partners, and in which males have larger testes in relation to their body weight, than in those species that mate with only one partner and have relatively small testes. Our results indicate that sexual selection by sperm competition has influenced the evolution of a specific component of male-gamete morphology, the volume of the sperm midpiece.

We classified different primate species as having either a single- or multiple-partner mating system on the basis of published descriptions of primary mating systems and behavioural information^{6,7}. We measured the lengths (in μm) and volumes (in μm^3) of sperm head, midpiece and flagellum in 31 species, representing 21 primate genera. We compared differences in sperm morphology between single- and multiple-

partner mating systems at both the genus and species levels, using comparative analysis of independent contrasts (CAIC)⁸. Critical α -levels were adjusted for multiple comparisons of sperm traits using the sequential Bonferroni technique.

Sperm lengths did not differ statistically between single- and multiple-partner mating systems. By contrast, sperm midpiece volumes differed significantly at both the genus ($F=20.86$, $P<0.01$) and species ($F=3.46$, $P<0.05$) levels (Fig. 1a). When sperm-head volumes associated with single- and multiple-partner mating systems were compared, there was no significant difference between genera or species ($44.01 \pm 2.71 \mu\text{m}^3$ versus $41.82 \pm 2.40 \mu\text{m}^3$, and $46.20 \pm 3.2 \mu\text{m}^3$ versus $44.56 \pm 2.89 \mu\text{m}^3$, respectively). The same was true for flagellum volumes ($9.74 \pm 0.82 \mu\text{m}^3$ versus $9.73 \pm 1.21 \mu\text{m}^3$, and $8.01 \pm 0.81 \mu\text{m}^3$ versus $9.69 \pm 0.88 \mu\text{m}^3$, respectively).

Because there is disagreement over the classification of primate mating systems, we also used relative testis weight as an indirect measure of likely sperm competition. We calculated residual values of testis weight from a regression plot of combined testes weight against body weight. We compared genera with relatively small testes (negative residuals) with those with relatively large

testes (positive residuals). Midpiece volumes were significantly greater in sperm from primate genera with relatively large testes (mean $\pm$ s.e.m.: $9.758 \pm 0.558 \mu\text{m}^3$ for genera with large testes; $5.360 \pm 0.609 \mu\text{m}^3$ for genera with small testes; $t = -4.57$, $P<0.001$). No significant difference was found when the analysis was repeated for sperm-head volume and flagellum volume.

We also tested for a possible correlation between relative testis weight and sperm midpiece volume, and found a strong effect ($r^2=0.49$, $P<0.0001$; Fig. 1b), although none of the other variables examined was correlated with relative testis weight. We conclude that sexual selection has probably influenced midpiece volume as well as testicular size.

Certain features of midpiece anatomy are highly heritable⁹. The volume of the midpiece is a more informative measure than its length, as this section contains a densely packed, helical array of mitochondria which provides energy for motility in the absence of glycolytic support^{10,11}. The volume is therefore a likely indicator of the degree of mitochondrial loading.

Mitochondrial loading also has a bearing on flagellar hydrodynamics^{10,11}, and the ATP content of the sperm of domestic fowl is positively correlated with its motility¹². The greater midpiece volumes we have observed in primates with multiple-partner mating systems, in which sperm competition is most prevalent, may therefore result from selection for increased mitochondrial loading and greater motility. This would mirror the situation in domestic fowl, in which high sperm motility is correlated with successful sperm competition¹³.

Matthew J. Anderson, Alan F. Dixon

Center for the Reproduction of Endangered Species,
Zoological Society of San Diego, PO Box 120551,
San Diego, California 92112-0551, USA

e-mail: manderson@sandiegozoo.org

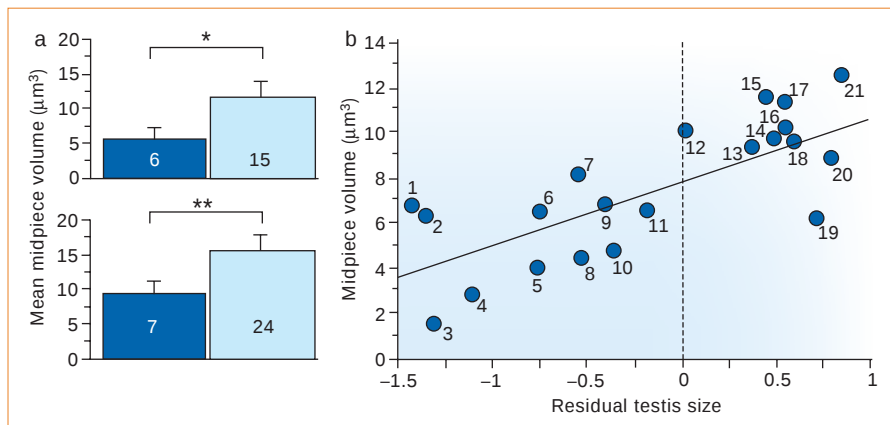


Figure 1 Sperm midpiece volume in relation to partner number and relative testis size. **a**, Comparison of sperm midpiece volume in primates with single-partner (black bars) or multiple-partner (white bars) mating systems at the genus level (top) and the species level (bottom; after comparative analysis of independent contrasts, CAIC). Numbers in bars show sample sizes (number of animals). Asterisk denotes $P<0.01$; double asterisk denotes $P<0.05$ (tablewide α -values maintained by sequential Bonferroni corrections). One hundred sperm from each of 106 specimens were measured at $\times 1,000$ magnification using a light-refractive microscope, a digital camera and Image NIH for Apple Macintosh. Midpiece volume was estimated using the formula for calculating the volume of a cylinder ($\pi \times$ midpiece width² $\times$ midpiece length). **b**, Midpiece volume (in μm^3) in relation to relative testis size in 21 primate genera ($y=2.55x+7.52$; $P<0.0001$). 1, *Gorilla*; 2, *Erythrocebus*; 3, *Callithrix*; 4, *Hylobates*; 5, *Pongo*; 6, *Theropithecus*; 7, *Saimiri*; 8, *Cebus*; 9, *Galago*; 10, *Homo*; 11, *Cercopithecus*; 12, *Lophocebus*; 13, *Mandrillus*; 14, *Pygathrix*; 15, *Lemur*; 16, *Eulemur*; 17, *Papio*; 18, *Nycticebus*; 19, *Microcebus*; 20, *Pan*; 21, *Macaca*.

1. Parker, G. A. in *Sperm Competition and the Evolution of Animal Mating Systems* (ed. Smith, R. L.) 1–60 (Academic, Orlando, 1984).
2. Birkhead, T. R. & Møller, A. P. *Sperm Competition and Sexual Selection* (Academic, London, 1998).
3. Short, R. V. *Adv. Study Behav.* **9**, 131–158 (1979).
4. Harcourt, A. H., Harvey, P. H., Larson, S. G. & Short, R. V. *Nature* **293**, 55–57 (1981).
5. Birkhead, T. R. *Promiscuity: An Evolutionary History of Sperm Competition and Sexual Conflict* (Faber & Faber, London, 2000).
6. Dixon, A. F. *Primate Sexuality: Comparative Studies of the Prosimians, Monkeys, Apes and Human Beings* (Oxford Univ. Press, New York, 1998).
7. Smuts, B. B., Cheney, D. L., Seyfarth, R. M., Wrangham, R. W. & Struhsaker, T. T. *Primate Societies* (Univ. Chicago Press, Chicago, 1987).
8. Purvis, A. & Rambaut, A. *Cabios* **11**, 247–251 (1995).
9. Woolley, D. M. & Beatty, R. A. *Nature* **215**, 94–95 (1967).
10. Cardullo, R. A. & Baltz, J. M. *Cell. Motil. Cytoskeleton* **19**, 180–188 (1991).
11. Katz, D. F. & Drobnis, E. Z. in *Fertilization in Mammals* (eds Bavister, B. D., Cummins, J. & Roldan, E. R. S.) 125–137 (Seroni Symp., Norwell, Massachusetts, 1990).
12. Froman, D. P. & Feltmann, A. J. *Biol. Reprod.* **58**, 379–384 (1998).
13. Birkhead, T. R., Martinez, J. G., Burke, T. & Froman, D. P. *Proc. R. Soc. Lond. B* **266**, 1759–1764 (1999).

Aerosol formation

Atmospheric particles from organic vapours

Aerosol particles produced over forested areas may affect climate by acting as nuclei for cloud condensation, but their composition (and hence the chemical species that drive their production) remains an open question. Here we show, to our knowledge for the first time, that these newly formed particles (3–5 nm in diameter) are composed primarily of organic species, such as *cis*-pinonic acid and pinic acid, produced by oxidation of terpenes in organic vapours released from the canopy^{1–4}.

Our technique combines aerosol electrical-mobility size measurements⁵ with organic (butanol) vapour growth rates in the cloud chamber of a modified condensation-particle counter⁶. Particles are introduced into a condensing flow-tube cloud chamber where they are subjected to a supersaturated

organic vapour. During transit through the flow tube, particles nucleate into organic cloud-droplets which are detected by light scattering on leaving the chamber.

Particles larger than 10 nm grow to the same final droplet size, whereas smaller particles grow less owing to the stronger Kelvin effect (lower effective supersaturation resulting from increased surface curvature at these sizes), resulting in a monotonic link between initial particle size and detected pulse height. In this size region, for a given size, the pulse is also dependent on particle chemical composition, owing to its solubility in the organic vapour. Therefore, if the initial size of the 3–10-nm particles is known, the composition can be determined as a function of growth in the organic vapour.

Figure 1a gives size-distribution measurements made before and during the initial stage of a nucleation event on 2 May 2000 at the Hyytiälä forest research station¹ over the boreal forest in Finland. During the event, a clear nucleation mode (3–6 nm)

is evident, together with a pre-existing particle mode at sizes over 30 nm.

Figure 1b shows the detected pulse height from droplet light scattering in the condensation-particle counter for the same periods. Before the event, the pulse height is typical of the pre-existing aerosol, with only a single distinct pulse height (and therefore droplet size) evident. During the event, additional smaller pulses are visible, corresponding to pulses generated by the 3–6-nm particles. Also shown are the calculated pulse-height responses, based on laboratory calibrations for the observed nucleation mode, which assume particle compositions of ammonium sulphate, pinic acid and *cis*-pinonic acid. The calculated spectra are generated by taking the laboratory pulse response for each size and composition and scaling it to the concentration measured in the corresponding mobility size range (5-nm particle laboratory pulses are shown as an example in Fig. 1b, inset).

A composition of ammonium sulphate for the measured nucleation-mode size distribution would not reproduce the observed pulse heights, but the observations agree well with those expected for pinic acid, and even better with those for *cis*-pinonic acid. Given that the growth of other common atmospheric aerosol inorganic species in the butanol vapour is very similar to ammonium sulphate, this nucleation mode cannot be inorganic in composition⁶.

Although our technique does not directly demonstrate that the condensing vapour responsible for producing the new particles is *cis*-pinonic or pinic acid, it reveals that an organic vapour with similar solubility in butanol must be responsible for the phenomenon. Further, given that these events occur in clean air where the ratio of biogenic to anthropogenic volatile organic compounds (VOCs) is high, and the fact that biogenic VOCs are considerably more reactive than anthropogenics⁷, it is unlikely that the particle production can be explained by anthropogenic species. Our results indicate that aerosol formation over forests is driven by condensable organic vapours.

Anthropogenic activities that result in increased concentrations of pollutants such as ozone will influence the conversion of natural VOCs into condensable vapours to generate natural aerosols. Complex feedback processes involving, for example, the coupling of emissions, radiative balance, and aerosol and cloud formation, possess uncertainties that must be determined if we are to predict future changes in global climate.

Colin D. O'Dowd*, **Pasi Aalto***,
Kaarle Hämeri†, **Markku Kulmala***,
Thorsten Hoffmann‡

*Division of Atmospheric Sciences, Department of Physical Sciences, PO Box 64, University of Helsinki, 00014 Helsinki, Finland

†Department of Experimental Physics, National University of Ireland, University Road,

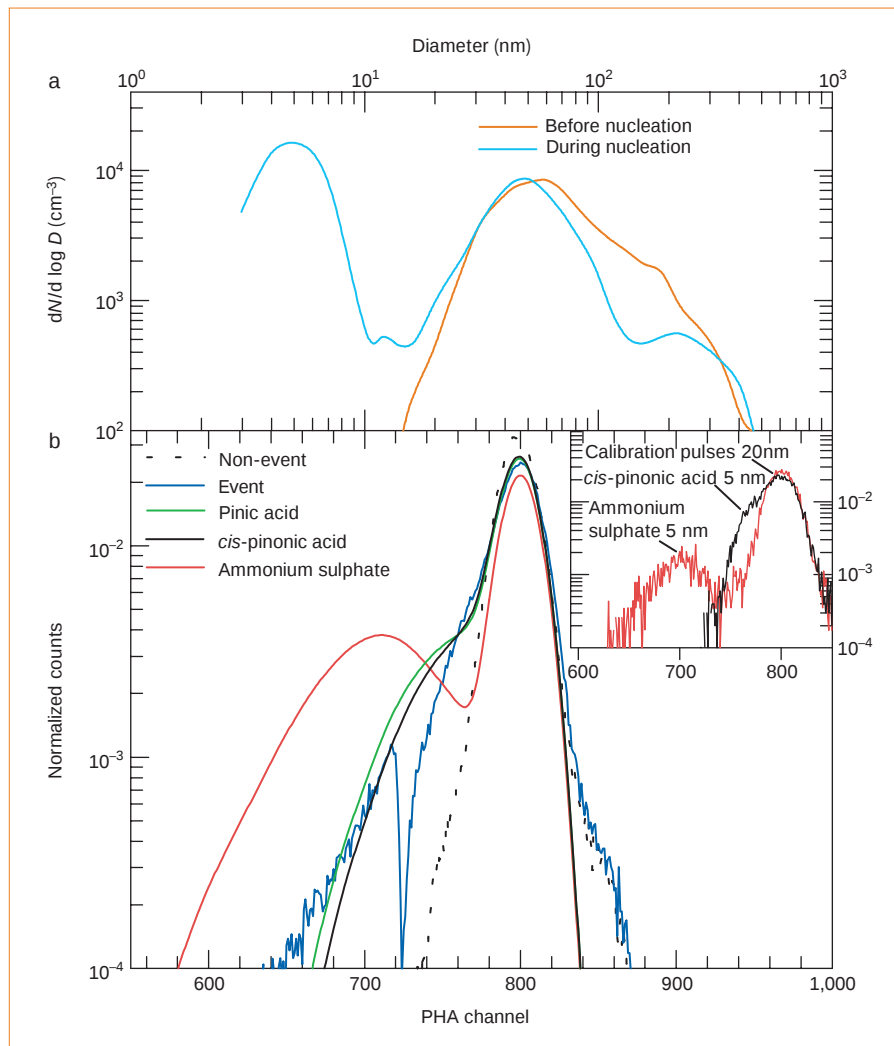


Figure 1 Size distribution of particles and their detection by a PHA (pulse-height analyser) condensation-particle counter. **a**, Particle size distributions taken before (04:00 local time) a nucleation event and during the initial stages of the nucleation event (11:30) on 2 May 2000 using a differential-mobility particle sizer. **b**, Pulse-height spectra measured before and during a nucleation burst. Also shown are calculated pulse-height spectra for the observed nucleation mode, based on laboratory calibrations for ammonium sulphate, pinic acid and *cis*-pinonic acid (for example, 5-nm calibration is shown in the inset, together with 20-nm particles for comparison).

Galway, Ireland

e-mail: colin.odowd@cmas.demon.co.uk

†Finnish Institute of Occupational Health,

Topeliuksenkatu 41 a A, FIN-00250 Helsinki, Finland

§Institute of Spectrochemistry and Applied Spectroscopy, 44139 Dortmund, Germany

1. Kulmala, M. et al. *Tellus* **53B**, 324–342 (2001).
2. Leitch, W. R. et al. *J. Geophys. Res.* **104**, 8095–8111 (1999).
3. Kavouras, I. G., Mihalopoulos, N. & Stephanou, E. G. *Nature* **395**, 683–686 (1998).
4. Birmili, W., Wiedensohler, A., Plass-Dülmer, C. & Berresheim, H. *Geophys. Res. Lett.* **27**, 2205–2208 (2000).
5. Mäkelä, J. M. et al. *Geophys. Res. Lett.* **24**, 1219–1222 (1997).
6. Marti, J. J., Weber, R. J., Saros, M. T. & McMurtry, P. H. *Aer. Sci. Technol.* **25**, 214–218 (1996).
7. Hoffmann, T. *EUROTRAC Newslett.* **21**, 12–21 (1999).

Competing financial interests: declared none.

Nanomechanics

Response of a strained semiconductor structure

The nanomechanical properties of thin silicon films will become increasingly critical in semiconductor devices, particularly in the context of substrates that consist of a silicon film on an insulating layer (known as silicon-on-insulator, or SOI, substrates). Here we use very small germanium crystals as a new type of nanomechanical stressor to demonstrate a surprising mechanical behaviour of the thin layer of silicon in SOI substrates, and to show that

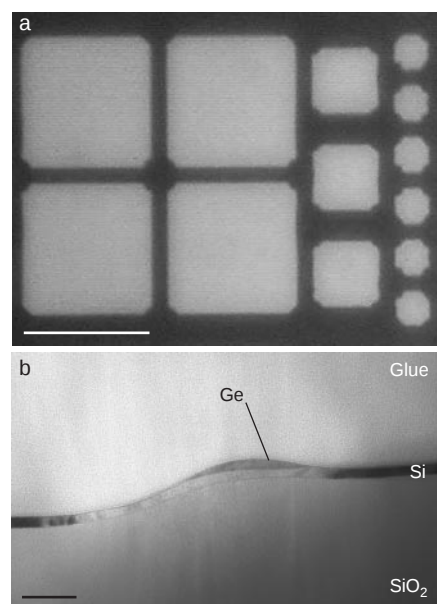


Figure 1 Anomalous local bending of the thin silicon-template layer in silicon-on-insulator (SOI) substrates induced by the growth of germanium (Ge) nanocrystals. **a**, Mesa structures patterned and etched on a SOI substrate wafer. **b**, Transmission electron microscope image showing a Ge nanocrystal and the bent Si-template layer underneath. The Ge nanocrystals, which grow in register with the Si lattice, have a 4% greater lattice constant than the Si, and thus cause it to bend locally. The local shear stress is sufficient to reduce the viscosity of the oxide underneath the thin Si layer so as to allow the oxide to flow locally. Scale bars, 20 μm (**a**) and 50 nm (**b**).

there is a large local reduction in the viscosity of the oxide on which the silicon layer rests. These findings have implications for the use of SOI substrates in nanoelectronic devices.

We use SOI substrates consisting of a handle wafer (a thick silicon (Si) layer), a thin oxide (a 400-nm thickness of SiO_2) and a very thin (10 nm) template layer of crystalline Si on top of the oxide. The template layer is patterned to form micrometre-sized (5–20 μm) mesas (Fig. 1a). About 10 monolayers of germanium (Ge; total thickness 1.6 nm) are deposited by molecular-beam epitaxy at 700 $^{\circ}\text{C}$. Germanium has a lattice constant 4% greater than that of silicon.

Figure 1b shows the formation of Ge nanocrystals (about 10 nm high, with 100-nm bases) that are crystallographically in register with the Si template, and an anomalous local bending of the Si template layer underneath each individual nanocrystal. The curvature underneath the islands is greater than 0.005 nm^{-1} . This new mode of local bending (Fig. 2a) of a nanometre-scale thin film is different from the commonly observed extended, uniform bending mode (Fig. 2b) that is induced by strained-layer film growth on thick Si(001) (refs 1, 2).

Our calculations show that the local bending curvature depends on the nanocrystal's density and shape (Fig. 2c, d). On a thick substrate, local bending is suppressed³, resulting in an overall extended bending that can be estimated using Stoney's formula⁴ and which is independent of nanocrystal density and shape, as would be the case in a uniform film of equivalent thickness.

The local bending mode and large bending magnitude indicate that the Si template layer behaves as a 'free-standing' layer during the growth of Ge nanocrystals, an outcome that can be achieved if SiO_2 acts as a fluid with substantial viscous flow. The viscosity of SiO_2 at 700 $^{\circ}\text{C}$ (the growth temperature) is ordinarily much too great for such a large degree of relaxation to occur. However, this viscosity can decrease almost exponentially with increasing applied shear stress⁵.

From the bending curvature, and hence the bending stress, we estimate⁵ that the viscosity of SiO_2 can be reduced by three to five orders of magnitude in the regions beneath the bent Si layer below the Ge nanocrystals. The relaxation time for SiO_2 flow is then reduced by a few orders of magnitude, to well within the deposition time of about 150 s. Thus, the large bending stress in the Si layer greatly enhances the viscous flow of SiO_2 , which in turn helps to increase the bending of the Si layer, because the Si film can then behave as a free-standing film.

The local stressor on the thin Si template layer of SOI substrate modifies both the mechanical properties of the Si layer and its electronic properties, providing a unique method for electronic (band) engineering on a nanometre scale. For these reasons,

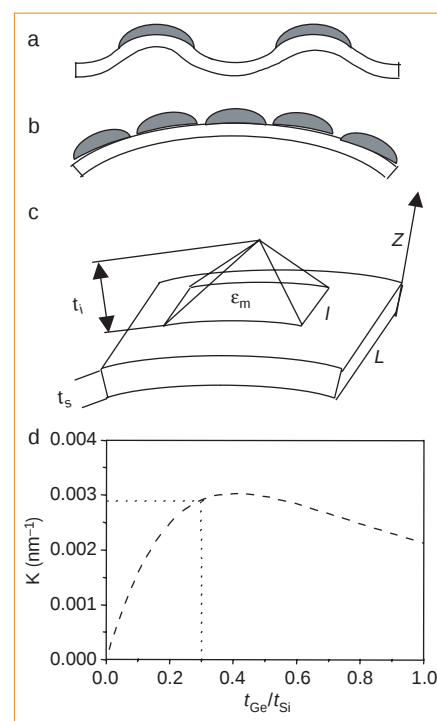


Figure 2 Different bending modes of thin silicon-template layer in SOI substrates induced by growing germanium nanocrystals. **a**, Local bending mode that occurs only when the density of nanocrystals (dotted domes) is low. **b**, Extended bending mode that occurs at high nanocrystal densities. For a thin Si layer, a transition from local to extended bending occurs with increasing nanocrystal density. **c**, Some parameters used to calculate the bending induced by a pyramidal nanocrystal acting as a local stressor. t_i is the thickness of the Ge nanocrystal, t_s is the thickness of the Si membrane, l is the base dimension of the nanocrystal, ϵ_m is the misfit strain in the nanocrystal, Z is the normal to the membrane, and L is the dimension of the bent region of membrane underneath the nanocrystal. **d**, The calculated bending curvature, K , of a free-standing, 10-nm Si layer as a function of the equivalent film thickness, t_{Ge} , normalized to the Si-layer thickness, t_{Si} . Arrow indicates the curvature (local thickness, local bending) that corresponds to the observed 10-nm Ge nanocrystal height; this is consistent with the data in Fig. 1b.

local stressors in SOI substrates could also become a significant issue for the semiconductor industry, which is increasingly using such substrates to manufacture devices.

Feng Liu*, **Paul Rugheimer†**, **E. Mateeva‡**, **D. E. Savage†**, **M. G. Lagally†**

*Department of Materials Science and Engineering, University of Utah, Salt Lake City, Utah 84112, USA

†Department of Materials Science and Engineering and Physics, University of Wisconsin, Madison, Wisconsin 53706, USA

e-mail: lagally@engr.wisc.edu

‡Colorado School of Mines, Golden, Colorado 80401, USA

1. Floro, J. A., Chason, E., Twisten, R. D., Hwang, R. Q. & Freund, L. B. *Phys. Rev. Lett.* **79**, 3946–3949 (1997).
2. Floro, J. A. et al. *Phys. Rev. B* **59**, 1990–1998 (1999).
3. Johnson, H. T. & Freund, L. B. *J. Appl. Phys.* **81**, 6081–6090 (1997).
4. Stoney, G. G. *Proc. R. Soc. Lond. A* **82**, 172–178 (1909).
5. Rafferty, C. S., Borucki, L. & Dutton, R. W. *Appl. Phys. Lett.* **54**, 1516–1518 (1989).

Competing financial interests: declared none.

An extensive network of coupling among gene expression machines

Tom Maniatis* & Robin Reed†

* Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA

† Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

Gene expression in eukaryotes requires several multi-component cellular machines. Each machine carries out a separate step in the gene expression pathway, which includes transcription, several pre-messenger RNA processing steps and the export of mature mRNA to the cytoplasm. Recent studies lead to the view that, in contrast to a simple linear assembly line, a complex and extensively coupled network has evolved to coordinate the activities of the gene expression machines. The extensive coupling is consistent with a model in which the machines are tethered to each other to form 'gene expression factories' that maximize the efficiency and specificity of each step in gene expression.

Eukaryotic gene expression is a complex stepwise process that begins with transcription initiation, elongation and termination (Fig. 1). During transcription, the nascent pre-mRNA is capped at the 5' end, introns are removed by splicing, and the 3' end is cleaved and polyadenylated. The mature mRNA is then released from the site of transcription and exported to the cytoplasm for translation. Superimposed on this pathway is an RNA surveillance system that eliminates aberrantly processed or mutant pre-mRNAs and mRNAs. Distinct machines carry out each of the steps in the gene expression pathway. Despite the unique reactions they catalyse, each machine also interfaces both physically and functionally with other machines in the pathway as detailed in several recent reviews^{1–5}. Here we discuss evidence that coupling is even more extensive than previously imagined. Indeed coupling occurs not only between sequential steps in the gene expression pathway but also between the earliest and latest steps (see Fig. 1).

In this review we address the reasons why such extensive coupling exists. As discussed below, coupling may solve many of the logistical problems inherent in the gene expression pathway. For example, the production of mature mRNA requires that the nascent pre-mRNA is sufficiently stable to complete its synthesis, processing and export. One of the many functions of the 5' cap is to protect the pre-mRNA from degradation. By tight coupling between the capping and transcription machineries rapid capping of the nascent pre-mRNA is ensured, thereby protecting it from degradation. Coupling also plays a critical role in gene expression by tethering machines to each other and to their substrates, a mechanism that dramatically increases the rate and specificity of enzymatic reactions. The possible consequences of tethering are illustrated by metazoan pre-mRNA splicing where small exons must be recognized in a vast sea of introns. This recognition problem may be solved at least in part by coupling transcription to splicing, which results in tethering splicing factors directly adjacent to the nascent pre-mRNA as it emerges from the polymerase. Tethering in general is widely used for regulating the activities within individual cellular machines⁶. As with the example of splicing above and as discussed throughout this review tethering is also used to coordinate activities between machines.

RNA polymerase II: coupling to RNA processing

Both the carboxy-terminal domain (CTD) of the RNA polymerase II large subunit and transcription elongation factors play central roles in coupling transcription to pre-mRNA processing. A number

of studies have shown that the CTD functions both as an assembly platform for and a regulator of transcription and pre-mRNA processing machines. The CTD consists of 27 heptad repeats in yeast and 52 in humans, and dynamic site-specific phosphorylation and dephosphorylation of these repeats is a critical mechanism for CTD function^{7–9}. Several different CTD kinases and at least one phosphatase have been implicated in this mechanism^{9–11}. The importance of the CTD in pre-mRNA processing is illustrated by the effects of certain CTD deletions which do not inactivate transcription but significantly decrease the efficiency of capping^{12,13}, splicing¹⁴ and polyadenylation¹⁴. The effects of these deletions on splicing and polyadenylation are not simply a secondary

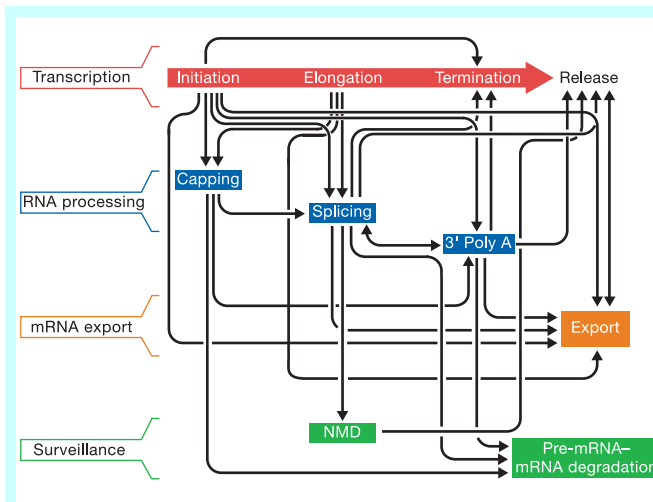


Figure 1 A complex network of coupled interactions in gene expression. The major steps of gene expression are indicated on the left and each stage in transcription is shown along the top in the red arrow. 'Release' indicates release of the mature mRNA from the site of transcription. The steps in RNA processing, mRNA export, nonsense-mediated decay (NMD) and RNA degradation are shown below the arrow. The black arrows indicate physical and/or functional coupling between two steps in gene expression. Each arrow is documented by published studies showing a physical and/or functional interaction between one or more components of gene expression machines. See text for a description of selected examples of coupling, and previous reviews for additional references and discussion^{2,3,5}.

consequence of the effects on capping¹⁵. Indeed recent studies indicate that the CTD independently promotes each step of the pre-mRNA processing pathway. Moreover distinct domains of the CTD appear to interact with factors required for different steps in pre-mRNA processing¹⁵, at least in humans.

Insight into the relationship between the CTD and the nascent pre-mRNA was provided by the recent determination of the three-dimensional structure of yeast RNA polymerase II (ref. 16). The 2.8 Å resolution structure revealed that the CTD is directly adjacent to the exit groove for the pre-mRNA. The CTD is 650 Å in length and is connected to the large subunit of RNA polymerase II through a flexible 280 Å linker. Thus the CTD appears to be of sufficient size and flexibility to interact with multiple components of the pre-mRNA processing machinery and to localize this machinery close to the pre-mRNA as it emerges from the exit groove of the polymerase.

The importance of positioning the CTD immediately adjacent to the nascent pre-mRNA is illustrated by the process of 5' end capping⁴. Capping takes place during the transition from transcription initiation to elongation when the nascent pre-mRNA is only 20–40 nucleotides long. During transcription initiation, serine 5 of the CTD heptad repeat is phosphorylated^{9,10}. This phosphorylation triggers a cascade of events beginning with the dissociation of transcription initiation factors from the CTD followed by the recruitment of the capping machinery and the allosteric activation of the capping enzyme^{8,17–19}. Serine 5 is subsequently dephosphorylated, resulting in release of the capping machinery^{8,18}. Serine 2 of the heptad repeat is then phosphorylated, leading to the recruitment of factors involved in subsequent steps of RNA processing^{8,9}. This switch in site-specific phosphorylation functions as a timing mechanism to coordinate the appearance of the 5' end of the pre-mRNA with capping^{8,9}.

In mammalian cells the transcription elongation factor P-TEFb is thought to be the serine 2 CTD kinase^{9,11,20}. In addition, P-TEFb phosphorylates another elongation factor, hSPT5 (refs 21 and 22). Another level of coupling between capping and transcription was revealed by the finding that hSPT5 interacts directly with and activates the capping enzyme²³. Thus P-TEFb plays a pivotal role in the temporal coordination of capping and transcriptional elongation⁴.

Pre-mRNA splicing is also coupled to transcription elongation in part through interactions between the splicing machinery and the transcription elongation factor TAT-SF1 (ref. 24). These interactions reveal yet another role for P-TEFb in coupling. Specifically, P-TEFb (which is recruited to promoters of cellular genes by transcription factors²⁵) recruits TAT-SF1²⁶, which in turn recruits splicing factors to the nascent pre-mRNA²⁴. *In vitro* studies reveal that the splicing machinery recruited by TAT-SF1 strongly stimulates transcription elongation²⁴. The role of TAT-SF1 appears to be conserved as the yeast TAT-SF1 homologue (CUS2) associates with splicing factors and regulates the assembly of splicing complexes²⁷. The connection between transcription elongation and splicing is further evidenced by *in vitro*²⁴ and *in vivo* studies showing that genes containing introns are more efficiently transcribed (reviewed in ref. 28). Finally, another transcription elongation factor, TFIIS, was recently shown to associate with a high molecular mass complex containing RNA polymerase II and several splicing factors²⁹. Thus, at least three elongation factors have been implicated in coupling transcription elongation to splicing: P-TEFb, TAT-SF1 and TFIIS.

As is the case with capping, transcription is also coupled to both splicing and polyadenylation through interactions between the CTD and the respective processing machineries^{2,14}. For example, a number of components of the polyadenylation machinery associate with the CTD² and these interactions are required for maximum levels of polyadenylation *in vivo*^{14,15}. Similarly a set of splicing-related proteins (SCAFs)³⁰ and various components of the splicing

machinery associate with the CTD^{1,2}. Among the CTD-associated splicing factors is Prp40 (ref. 31) which is thought to function in bringing the 5' and 3' splice sites together during the splicing reaction³².

Coupling between the splicing machinery and the transcription initiation and elongation complexes may solve long-standing problems in metazoan pre-mRNA splicing. Most metazoan pre-mRNAs contain very short exons (~140 nucleotides)³³ whereas the introns are often tens of thousands of nucleotides long. Indeed one intron in human neurexin pre-mRNA is approximately 480,000 nucleotides in length (B. Graveley, personal communication). Thus several hours pass between the time of synthesis of the 5' and 3' splice sites of this intron. In addition, metazoan pre-mRNAs contain multiple exons and introns, and splice-site sequences are highly degenerate. Thus the splicing machinery must specifically recognize each exon in the context of enormous introns and bring them together in the correct 5' to 3' order.

Each of these problems may be solved by tethering the splicing machinery to both the CTD and the transcription elongation complex. This tethering concentrates the splicing machinery directly adjacent to the nascent pre-mRNA as it emerges from the polymerase^{2,24,34}. The specific recognition of exons may be facilitated by this tethering, which effectively reduces the kinetics of the interaction between the splicing machinery and the pre-mRNA from first to zero order. With first-order kinetics, factors interact by free diffusion, dramatically decreasing the probability of an interaction. Consequently by coupling transcription to splicing, the splicing machinery has a strong competitive advantage over the abundant non-specific RNA-binding proteins that can contact the pre-mRNA only through free diffusion. As a result of coupling, the splicing machinery has an increased probability of interacting with the pre-mRNA, thereby increasing the specific recognition of exon sequences. Coupling between the transcription and splicing machineries is also likely to be important in ensuring that exons are joined in the correct 5' to 3' order during splicing. This order may be established by tethering each newly synthesized exon and the adjacent splice sites to the CTD until the next exon emerges from the exit groove of the polymerase^{24,34}.

Transcription factors and coactivators: coupling to splicing

Further evidence that transcription is coupled to splicing is provided by recent reports of physical and in some cases functional interactions between the preinitiation complex and the SR protein family of splicing factors^{35–38}. This family, which contains arginine/serine (RS) domains³⁹, binds to specific exon sequences known as splicing enhancers and recruits the splicing machinery to the 5' and 3' splice sites^{39,40}. Thus SR proteins play a critical role in recognizing exons for inclusion in the spliced mRNA⁴⁰. An example of an SR-like protein thought to have a role in coupling transcription to splicing is the transcriptional coactivator PGC-1. This coactivator, which contains both an RS domain and an RNA-binding domain, is recruited to promoters activated by PPAR γ and other nuclear receptors⁴¹. Neither domain appears to be required for transcription but both are required for the accumulation of mRNAs transcribed from PGC-1-dependent promoters³⁷. The PGC-1 protein colocalizes with splicing factors in the nucleus and interacts directly with SR proteins *in vitro*. This and related examples^{35,36,38} suggest a common theme for transcription-coupled splicing in which transcription factors recruit SR proteins to the transcription machinery. The proximity of these proteins to the nascent pre-mRNA during transcription may in turn promote splicing.

Coupling of transcription to splicing was also revealed in studies showing that transcription of pre-mRNA by different promoters can generate different alternatively spliced mRNAs^{42,43}. Two models which are not mutually exclusive have been proposed to explain this promoter-specific alternative splicing. In the first, particular

SR protein family members are differentially recruited to promoters and are then handed off to cognate splicing enhancers to promote inclusion of specific exons in the mRNA⁴³. In the second, the pattern of splicing is determined by the rate of transcription elongation which is dictated in turn by the particular promoter^{34,44}. In the second model different pre-mRNA secondary structures would be generated depending on the rate of transcription. The alternate structures could determine the availability of sequence elements recognized by the splicing machinery. This model is consistent with the observation that the pattern of alternative splicing can differ depending on whether the splicing machinery acts on a fully synthesized pre-mRNA, or whether splicing occurs co-transcriptionally^{34,45}. In addition, transcriptional pausing or conditions that can decrease the rate of elongation can affect the pattern of alternative splicing^{34,44,46}.

Coupling capping, splicing and polyadenylation

As proposed for coupling between transcription and pre-mRNA processing, coupling among all of the different pre-mRNA processing steps also seems to be critical for gene expression. For example, proteins that bind to the 5' cap of pre-mRNAs interact with splicing factors and promote recognition of the cap-proximal 5' splice site^{47,48}. Splicing factors that associate with the terminal 3' intron also interact with downstream polyadenylation factors and promote 3' end cleavage and polyadenylation^{49,50}. Reciprocally, polyadenylation complexes can promote splicing of the upstream 3' terminal intron³³. These interactions, in conjunction with the recruitment of the splicing machinery to internal exons in the pre-mRNA⁵¹, play an important role in ensuring that the correct splice sites are recognized and engaged in the splicing reaction, and that capping and polyadenylation are timely, accurate and efficient.

Coupling the earliest and latest steps of gene expression

Polyadenylation, transcription termination (release of RNA polymerase from the DNA) and release of the RNA from the site of transcription are the final steps in the production of mature mRNA. A premature event in any of these steps would result in the release of an incompletely synthesized or partially processed pre-mRNA. This problem seems to be circumvented by extensive coupling among the late steps in pre-mRNA synthesis. For example, both a functional polyadenylation signal and polyadenylation factors are required for normal transcription termination^{52,53}. Moreover RNA polymerase II CTD truncations that are defective in polyadenylation are likewise defective in termination^{14,54}. A number of models for coupling polyadenylation and termination have been proposed but a general view has not emerged, owing at least in part to differences in assay systems^{2,52,55–57}.

Recent studies have shown that transcription initiation and termination may also be coupled and a polyadenylation factor may mediate this coupling^{58,59}. The human transcriptional coactivator protein PC4, which interacts directly with the polyadenylation factor CstF, provides an example. The yeast homologues of these proteins interact genetically, and both are required for efficient transcription termination⁵⁹. On the basis of these observations, it was proposed that PC4 functions as an anti-terminator by inhibiting the termination activity of CstF during transcription elongation. PC4 and CstF would then dissociate from each other at the polyadenylation site, allowing the CstF to function in both polyadenylation and termination⁵⁹. This speculative model is attractive because events that occur during the earliest steps of gene expression would ensure the correct temporal order of downstream steps in the pathway.

Subsequent to transcription termination, the mature mRNA is released from the site of synthesis and transported to nuclear pores. Even this release step is coupled to both upstream and downstream events⁶⁰. For example, unlike wild-type pre-mRNAs, pre-mRNAs containing splicing mutations accumulate at the site of transcrip-

tion. The mutant pre-mRNAs may form abortive splicing complexes that fail to release normally or splicing itself may be required for release. The coupling of release and mRNA export is also suggested by the observation that RNA transcripts accumulate at the site of synthesis in a number of mRNA export mutants⁶¹. Whether or not this is a direct effect of coupling remains to be established.

Coupling and the gene expression factory model

The physical coupling between transcription and pre-mRNA processing components may also play an important role in the assembly, storage, and intranuclear transport of different gene expression machines. Transcription and splicing factors are concentrated in two types of nuclear regions known as Cajal bodies and nuclear speckles^{62,63}. Time course studies suggest that transcription and pre-mRNA processing components newly imported into the nucleus first localize to Cajal bodies, and are ultimately stored in the nuclear speckles^{62,64}. To activate gene expression, transcription and pre-mRNA processing components move from speckles to the gene^{65–67}. As with several of the coupling steps mentioned above this intranuclear transport of splicing components requires an intact CTD⁶⁷. Although the mechanism of transport is not understood it appears that phosphorylation of SR proteins is required for their release from speckles⁶⁶, and the movement of SR proteins within the nucleus is rapid and ATP-independent⁶⁸. Thus SR proteins diffuse freely in the nucleus, rapidly moving from storage depots to their sites of action. Coupling between machines in such an environment could have an essential role in coordinating their assembly and localization.

An intriguing question is whether the transcription machinery moves along the DNA carrying all of the coupled machines with it or is anchored in an immobilized 'gene expression factory'^{14,69} and reels in the template DNA as the pre-mRNA is continuously extruded in a highly localized manner⁷⁰. At present it is not possible to distinguish rigorously between these two models. However the immobilized gene expression factor model⁷⁰ has considerable appeal from the perspective of the extensive coupling in gene expression. Consistent with this model, there is evidence that both the transcription and splicing machineries associate with the nuclear substructure, and that newly synthesized pre-mRNA is concentrated in discrete nuclear foci^{63,70}. These foci may correspond to the gene expression factories, containing all of the tightly coupled machines required to produce mature mRNA. The CTD would then function as the linchpin of these factories, playing a central role in the assembly and coordination of all of the machines.

A working model for coupling transcription and pre-mRNA processing based on the gene expression factory model is shown in Fig. 2. In this model capping (CAP), splicing (SF) and polyadenylation (pA) factors are recruited to the transcription preinitiation complex (PIC) (see Fig. 2a). At this stage transcription factors (TF) are bound to the unphosphorylated CTD. Upon transcription initiation the CTD is phosphorylated, the PIC is released and capping, splicing and polyadenylation components associate with the CTD (Fig. 2b). Shortly after pre-mRNA synthesis begins, the 5' end is capped, and the capping machinery then dissociates from the CTD (Fig. 2b). As the DNA is reeled into the gene expression factory, and the nascent pre-mRNA is extruded through the exit channel of the RNA polymerase, splicing factors are systematically transferred from the CTD to the exons (Fig. 2c), where they recruit the rest of the splicing machinery. Splicing factors are also transferred via the elongation complex (not shown). The interaction between splicing factors bound to the CTD with those bound to the exon functions to tether the exon to the CTD until the next exon emerges from the exit pore of the polymerase (Fig. 2d). This tethering plays a role in recognizing exons, and in ensuring the correct 5' and 3' association of sequentially synthesized exons. The splicing factors transferred from the CTD to the nascent pre-mRNA

are rapidly replaced by new splicing factors which are optimally positioned for the appearance of the next exon. This reloading mechanism is consistent with electron tomography studies of Balbiani ring chromosomes showing that splicing factors are continuously added to and released from nascent transcripts during

transcription elongation⁷¹. The splicing machinery associates with the exons in a 5' to 3' order as they are synthesized. However as indicated in the figure the covalent joining of exons does not necessarily proceed in a 5' to 3' direction owing to inherent differences in rates of splicing catalysis between each exon pair.

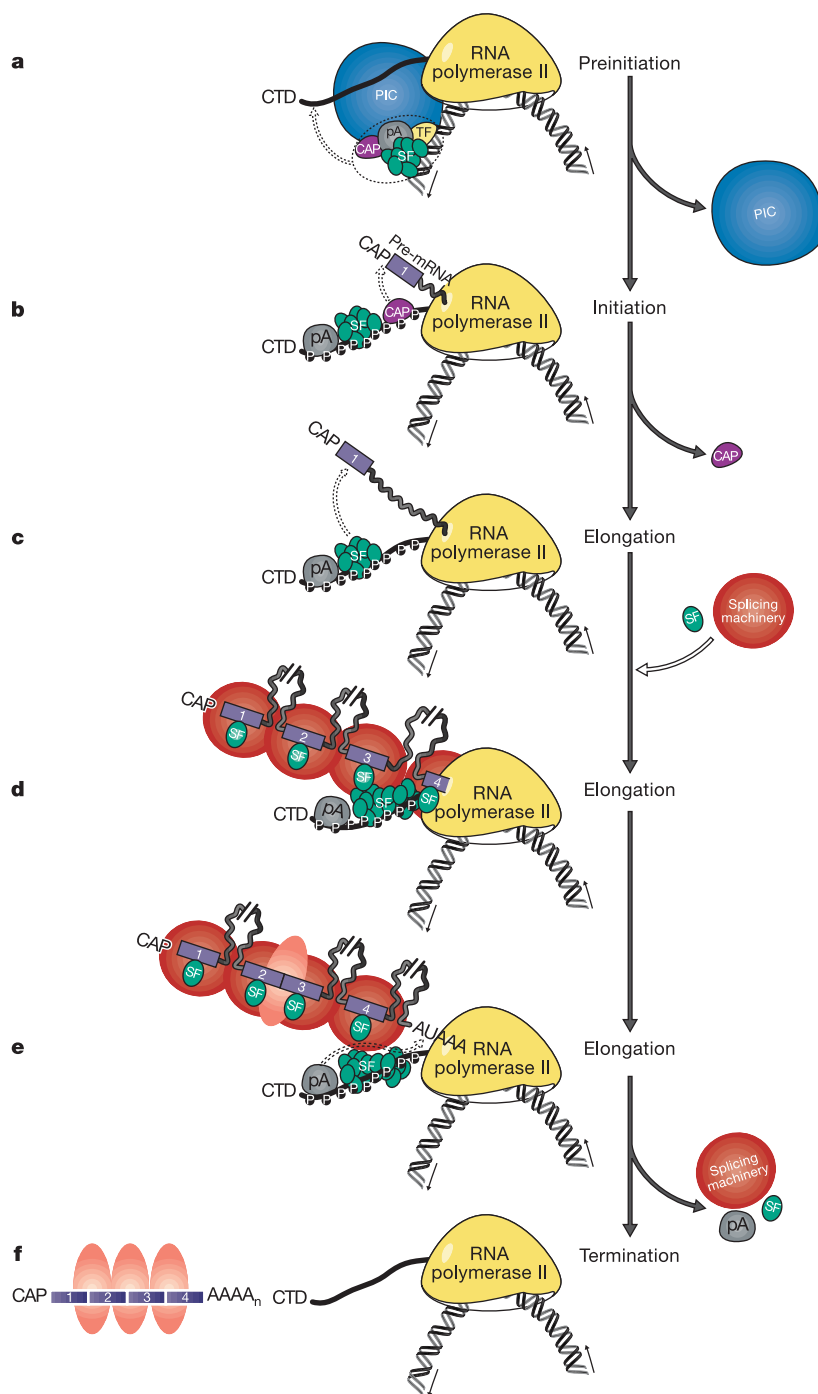


Figure 2 Gene expression factory model for coupling steps in gene expression. In this model the gene expression factory is anchored to the nuclear substructure and the DNA is reeled through the RNA polymerase as the nascent RNA is extruded through its exit channel. The machineries involved in transcription, capping, splicing and polyadenylation are shown. The shaded pink ovals over the spliced exons represent the

mRNA complexes formed near the exon–exon junctions during the splicing reaction. See text for details of parts **a–f**. PIC, preinitiation complex; TF, transcription factors; CTD, carboxy-terminal domain; CAP, capping factor; SF, splicing factor; pA, polyadenylation factor; P, phosphorylated CTD.

Splicing leads to the formation of a specific complex of proteins on the spliced exons (messenger RNA–ribonucleoprotein (mRNP) complex, described below). When the polymerase approaches the end of the transcript, the polyadenylation factors function (Fig. 2e). Finally the CTD is dephosphorylated, the processing factors are released and the mature mRNP is released from the site of transcription (Fig. 2f).

Coupling splicing to mRNA export

Splicing occurs at the site of pre-mRNA synthesis before the mature mRNA is released for nuclear export. Despite the different nuclear locations of splicing and export, recent studies indicate that these two processes are directly coupled. Evidence for this coupling initially came from the observation that mRNAs generated by splicing are more efficiently exported than their identical counterparts transcribed from a complementary DNA⁷². This effect of splicing on export was explained by the finding that spliced mRNAs (but not cDNA transcripts) are assembled into a distinct mRNP complex that promotes efficient export⁷². This complex, or at least certain components of it, was subsequently shown to assemble adjacent to newly formed exon–exon junctions⁷³. The increased export efficiency of the spliced mRNP is thought to be due to recruitment of the mRNA export factor ALY to the mRNA during the splicing reaction^{74,75}. The splicing factor UAP56, which interacts directly with ALY, plays a role in recruiting ALY to the spliced mRNA^{76–78}. The mRNP complex also contains a number of other factors^{74,75,79–83}, possibly including TAP⁷⁹, a protein that associates with the nuclear pore and is thought to be an mRNA export receptor^{84,85}. Studies in yeast indicate that both TAP (Mex67) and UAP56 (Sub2) interact directly with the same region on ALY (Yra1)⁷⁷. This observation together with other work suggests a model in which ALY is first recruited to the mRNA by interactions with UAP56 during the splicing reaction^{76–78,86}. In a subsequent step, a hand-off occurs in which the ALY/TAP interaction is established, thus delivering the mRNP to the nuclear pore for export⁷⁷.

As mentioned above, metazoan introns are usually large, whereas exons are minute by comparison. Thus, specific mechanisms must exist to ensure that the vast amount of intron sequences are retained in the nucleus and degraded, while the relatively small amount of mature mRNA is exported. The coupling between splicing and export may provide another mechanism for this discrimination⁸⁷. Specifically the mRNA is packaged into a splicing-dependent complex containing export factors while the large intron sequences are packaged into hnRNP complexes, which may play a role in nuclear retention. Thus, the coupling between splicing and mRNA export may provide a critical checkpoint for proper gene expression⁸⁷.

A small number of metazoan and most yeast pre-mRNAs lack introns. It is nevertheless possible that the same conserved mRNA export machinery described above is involved in intronless mRNA export, but is recruited by other mechanisms. In support of this notion, UAP56 is required for export of both spliced mRNAs and intronless mRNAs^{76–78,86}. In metazoan intronless mRNAs, specific mRNA sequence elements are required for export. In one such mRNA, this element associates with members of the SR family of splicing factors, which are thought to mediate export of the intronless mRNA⁸⁸. The SR proteins may either recruit the conserved export machinery or play a direct role in export themselves. In both yeast and metazoans the export of intronless mRNAs may also be coupled to polyadenylation⁸⁹.

The coupling between splicing and export provides an example of linking two machineries with spatially distinct locations, one at the site of transcription and the other at the nuclear pore. This long-range coupling is further illustrated by a recent study in yeast indicating that mRNA export is also coupled to transcription⁹⁰. In this case Np13, an SR-like protein thought to function in export,

was shown to interact genetically and physically with components of the transcription machinery. In addition, both Np13 and the mRNA export factor Yra1 (that is, yeast ALY) are thought to associate with the pre-mRNA during its synthesis⁹⁰. Another example of long-range coupling is provided by the Hrp1 protein, which may associate with pre-mRNA near or at the site of transcription initiation⁸. This protein is also required for polyadenylation⁹¹, mRNA export⁹², and nonsense-mediated decay (NMD, see below)⁹³.

The coupling of export to both transcription and pre-mRNA processing requires that export machinery components bind to the nascent pre-mRNA at its site of synthesis and processing, and then remain bound until the mature mRNA reaches the nuclear pores. These interactions have several interesting implications. First, the proteins involved must form highly stable complexes with the pre-mRNA in order to remain bound throughout the entire pathway. Second, these proteins must also have high affinity for their ultimate targets at the nuclear pores so that they are recruited efficiently. Finally, mRNA export proteins bound to pre-mRNA so early in the gene expression pathway must be prevented from functioning at the incorrect time and exporting nascent pre-mRNA or pre-mRNA that is not fully processed. The binding of late-acting proteins at such early steps in gene expression would increase the efficiency and specificity of gene expression.

Coupling pre-mRNA processing to nonsense-mediated decay

During the course of mRNA production, specific RNA surveillance mechanisms operate to detect and eliminate defective pre-mRNA/mRNA⁹⁴. The best characterized of these mechanisms is NMD, in which pre-mRNAs containing premature termination codons are targeted for degradation. This process was first linked to splicing by the observation that pre-mRNAs containing stop codons located greater than 50 nucleotides (nts) upstream from an intron are targeted for destruction⁹⁴. This finding implicated the exon junction and hence splicing in NMD. The recent identification of the spliced mRNP complex⁷², and the observation that it forms near exon–exon junctions^{73,75,79,81,83}, have provided new insights into the mechanism for coupling splicing to NMD. These studies revealed that in addition to the mRNA export factors ALY and TAP, the exon junction complex contains factors required for NMD^{73,75,79,80,82,83}, including hUPF3, the human homologue of an essential yeast NMD protein^{94–96}. The hUPF3 protein is recruited to the exon junction complex by the splicing factor RNPS1 (ref. 81), and another protein called Y14 (ref. 83). Importantly, these interactions persist in the cytoplasm, thereby establishing a direct link between pre-mRNA splicing in the nucleus and NMD in the cytoplasm^{79,81,82}.

The exon junction-marking proteins (for example, Y14, RNPS1 and UPF3) play a critical role in the cytoplasmic degradation of transcripts containing a premature termination codon. These proteins, which are normally stripped from the mRNA during the first round of translation, instead remain bound because the movement of the ribosome is stopped by the premature termination codon^{79,81,82}. The presence of these proteins then leads to the recruitment of factors that trigger rapid mRNA decay. Thus, it has been proposed that coupling between the splicing and NMD machineries carries a molecular memory of the intron to the cytoplasm^{79,81,82}. A model for coupling splicing, mRNA export and NMD is shown in Fig. 3.

In addition to NMD in the cytoplasm, many pre-mRNAs containing nonsense mutations are eliminated in the nucleus before they can be exported to the cytoplasm⁹⁷. Although the mechanism for this nuclear NMD is not well understood it appears that NMD is coupled to the release of the nascent pre-mRNA from the site of transcription⁹⁸. Moreover, mRNAs containing premature termination codons may be subject to surveillance as they move from the nucleus to the cytoplasm⁹⁹.

In yeast, where most genes lack introns, specific RNA sequence

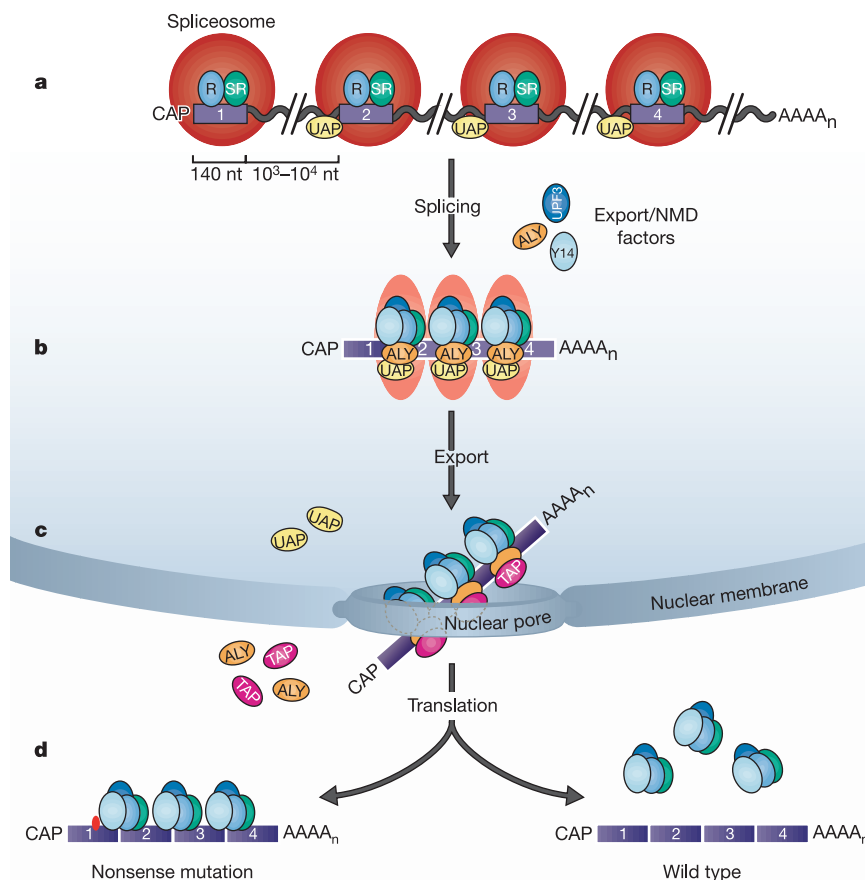


Figure 3 Model for coupling splicing to mRNA export and nonsense-mediated decay. In this model the splicing factor UAP56 (UAP, in yellow) associates with the pre-mRNA during spliceosome assembly along with other splicing factors including RNPS1 (R) and other SR proteins (a). UAP56 then recruits the mRNA export factor ALY along with NMD components including Y14 and UPF3 (b). All of these proteins form a complex that associates with the RNA near the exon–exon junction. UAP56 may then dissociate from ALY and be replaced by the export factor TAP, which mediates export of the mRNA

through the nuclear pore complex (c). In the cytoplasm the export machinery is released and the NMD machinery remains associated with the exon–exon junctions. When the mRNA is wild type, the NMD machinery is displaced by the translating ribosome. In contrast, in the presence of a premature stop codon (represented by the red dot) (d) upstream from an exon–exon junction translation is arrested and the NMD machinery is not released but instead recruits additional components that trigger rapid degradation of the mRNA. nt, nucleotides.

elements called DSEs (downstream sequence elements) are required for NMD¹⁰⁰. DSEs, which are located within protein coding sequences, act on upstream premature termination codons. Recently, Hrp1, the same protein implicated in both polyadenylation and mRNA export, was shown to bind specifically to DSEs, and to be required for NMD⁹³. In addition, Hrp1 specifically interacts with other proteins required for NMD. Thus, similarly to the EJC complex, which marks intron-containing pre-mRNAs for export and NMD, the Hrp1 protein appears to form an export/NMD complex on DSEs⁹³. Polyadenylation, mRNA export and NMD are therefore coupled through the multiple activities of the Hrp1 protein.

Discussion

The vast network of coupling discussed here and illustrated in Fig. 1 reveals that virtually every step in gene expression, from the earliest to the latest, is coupled. The emerging picture is that gene expression is carried out in immobilized gene expression factories consisting of a large number of interacting machines that orchestrate the multiple steps in the gene expression pathway. Understanding the inner workings of these factories will reveal an even greater

number of interactions between the components of gene expression machines. ☐

- Bentley, D. Coupling RNA polymerase II transcription with pre-mRNA processing. *Curr. Opin. Cell Biol.* **11**, 347–351 (1999).
- Hirose, Y. & Manley, J. L. RNA polymerase II and the integration of nuclear events. *Genes Dev.* **14**, 1415–1429 (2000).
- Proudfoot, N. Connecting transcription to messenger RNA processing. *Trends Biochem. Sci.* **25**, 290–293 (2000).
- Shatkin, A. J. & Manley, J. L. The ends of the affair: capping and polyadenylation. *Nature Struct. Biol.* **7**, 838–842 (2000).
- Cramer, P. *et al.* Coordination between transcription and pre-mRNA processing. *FEBS Lett.* **498**, 179–182 (2001).
- Ptashne, M. & Gann, A. *Genes and Signals* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2002).
- Dahmus, M. E. Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**, 19009–19012 (1996).
- Komarnitsky, P., Cho, E. J. & Buratowski, S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460 (2000).
- Cho, E. J., Kobor, M. S., Kim, M., Greenblatt, J. & Buratowski, S. Opposing effects of Ctk1 kinase and Fcp1 phosphatase at Ser 2 of the RNA polymerase II C-terminal domain. *Genes Dev.* **15**, 3319–3329 (2001).
- Trigon, S. *et al.* Characterization of the residues phosphorylated *in vitro* by different C-terminal domain kinases. *J. Biol. Chem.* **273**, 6769–6775 (1998).
- Price, D. H. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.* **20**, 2629–2634 (2000).

12. McCracken, S. *et al.* 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306–3318 (1997).
13. Cho, E. J., Takagi, T., Moore, C. R. & Buratowski, S. mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**, 3319–3326 (1997).
14. McCracken, S. *et al.* The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* **385**, 357–361 (1997).
15. Fong, N. & Bentley, D. L. Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev.* **15**, 1783–1795 (2001).
16. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**, 1863–1876 (2001).
17. Ho, C. K. & Shuman, S. Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* **3**, 405–411 (1999).
18. Schroeder, S. C., Schwer, B., Shuman, S. & Bentley, D. Dynamic association of capping enzymes with transcribing RNA polymerase II. *Genes Dev.* **14**, 2435–2440 (2000).
19. Cho, E. J., Rodriguez, C. R., Takagi, T. & Buratowski, S. Allosteric interactions between capping enzyme subunits and the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**, 3482–3487 (1998).
20. Taube, R., Lin, X., Irwin, D., Fujinaga, K. & Peterlin, B. M. Interaction between P-TEFb and the C-terminal domain of RNA polymerase II activates transcriptional elongation from sites upstream or downstream of target genes. *Mol. Cell Biol.* **22**, 321–331 (2002).
21. Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D. & Handa, H. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription *in vitro*. *EMBO J.* **17**, 7395–7403 (1998).
22. Kim, J. B. & Sharp, P. A. Positive transcription elongation factor B phosphorylates hSPT5 and RNA polymerase II carboxyl-terminal domain independently of cyclin-dependent kinase-activating kinase. *J. Biol. Chem.* **276**, 12317–12323 (2001).
23. Wen, Y. & Shatkin, A. J. Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev.* **13**, 1774–1779 (1999).
24. Fong, Y. W. & Zhou, Q. Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–933 (2001).
25. Barboric, M., Nissen, R. M., Kanazawa, S., Jabrane-Ferrat, N. & Peterlin, B. M. NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* **8**, 327–337 (2001).
26. Kim, J. B., Yamaguchi, Y., Wada, T., Handa, H. & Sharp, P. A. Tat-SF1 protein associates with RAP30 and human SPT5 proteins. *Mol. Cell Biol.* **19**, 5960–5968 (1999).
27. Yan, D. *et al.* CUS2, a yeast homolog of human Tat-SF1, rescues function of misfolded U2 through an unusual RNA recognition motif. *Mol. Cell Biol.* **18**, 5000–5009 (1998).
28. Ares, M. Jr, Grate, L. & Pauling, M. H. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA* **5**, 1138–1139 (1999).
29. Robert, F., Blanchette, M., Maes, O., Chabot, B. & Coulombe, B. A human RNA polymerase II-containing complex associated with factors necessary for spliceosome assembly. *J. Biol. Chem.* **277**, 9302–9306 (2002).
30. Corden, J. L. & Patturajan, M. A CTD function linking transcription to splicing. *Trends Biochem. Sci.* **22**, 413–416 (1997).
31. Morris, D. P. & Greenleaf, A. L. The splicing factor, Prp40, binds the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **275**, 39935–39943 (2000).
32. Abovich, N. & Rosbash, M. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89**, 403–412 (1997).
33. Berget, S. M. Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**, 2411–2414 (1995).
34. Goldstrohm, A. C., Greenleaf, A. L. & Garcia-Blanco, M. A. Co-transcriptional splicing of pre-messenger RNAs: considerations for the mechanism of alternative splicing. *Gene* **277**, 31–47 (2001).
35. Lai, M. C., Teh, B. H. & Tarn, W. Y. A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. *J. Biol. Chem.* **274**, 11832–11841 (1999).
36. Ge, H., Si, Y. & Wolffe, A. P. A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. *Mol. Cell* **2**, 751–759 (1998).
37. Monsalve, M. *et al.* Direct coupling of transcription and mRNA processing through the thermogenic coactivator PGC-1. *Mol. Cell* **6**, 307–316 (2000).
38. Martinez, E. *et al.* Human staga complex is a chromatin-acetylating transcription coactivator that interacts with pre-mRNA splicing and DNA damage-binding factors *in vivo*. *Mol. Cell Biol.* **21**, 6782–6795 (2001).
39. Graveley, B. R. Sorting out the complexity of SR protein functions. *RNA* **6**, 1197–1211 (2000).
40. Graveley, B. R., Hertel, K. J. & Maniatis, T. SR proteins are 'locators' of the RNA splicing machinery. *Curr. Biol.* **9**, R6–R7 (1999).
41. Puigserver, P. *et al.* A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
42. Cramer, P., Pesce, C. G., Baralle, F. E. & Kornblihtt, A. R. Functional association between promoter structure and transcript alternative splicing. *Proc. Natl Acad. Sci. USA* **94**, 11456–11460 (1997).
43. Cramer, P. *et al.* Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol. Cell* **4**, 251–258 (1999).
44. Kadener, S. *et al.* Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. *EMBO J.* **20**, 5759–5768 (2001).
45. Eperon, L. P., Graham, I. R., Griffiths, A. D. & Eperon, I. C. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* **54**, 393–401 (1988).
46. Roberts, G. C., Gooding, C., Mak, H. Y., Proudfoot, N. J. & Smith, C. W. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res.* **26**, 5568–5572 (1998).
47. Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C. & Mattaj, J. W. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* **10**, 1683–1698 (1996).
48. Colot, H. V., Stutz, F. & Rosbash, M. The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.* **10**, 1699–1708 (1996).
49. Vagner, S., Vagner, C. & Mattaj, J. W. The carboxyl terminus of vertebrate poly(A) polymerase interacts with U2AF 65 to couple 3'-end processing and splicing. *Genes Dev.* **14**, 403–413 (2000).
50. McCracken, S., Lambermon, M. & Blencowe, B. J. SRm160 splicing coactivator promotes transcript 3'-end cleavage. *Mol. Cell Biol.* **22**, 148–160 (2002).
51. Reed, R. Mechanisms of fidelity in pre-mRNA splicing. *Curr. Opin. Cell Biol.* **12**, 340–345 (2000).
52. Proudfoot, N. J. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14**, 105–110 (1989).
53. Birse, C. E., Minvielle-Sebastia, L., Lee, B. A., Keller, W. & Proudfoot, N. J. Coupling termination of transcription to messenger RNA maturation in yeast. *Science* **280**, 298–301 (1998).
54. McCracken, S. *et al.* Role of RNA polymerase II carboxy-terminal domain in coordinating transcription with RNA processing. *Cold Spring Harbor Symp. Quant. Biol.* **63**, 301–309 (1998).
55. Dye, M. J. & Proudfoot, N. J. Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. *Cell* **105**, 669–681 (2001).
56. Osheim, Y. N., Proudfoot, N. J. & Beyer, A. L. EM visualization of transcription by RNA polymerase II: downstream termination requires a poly(A) signal but not transcript cleavage. *Mol. Cell* **3**, 379–387 (1999).
57. Tran, D. P., Kim, S. J., Park, N. J., Jew, T. M. & Martinson, H. G. Mechanism of poly(A) signal transduction to RNA polymerase II *in vitro*. *Mol. Cell Biol.* **21**, 7495–7508 (2001).
58. Aranda, A. & Proudfoot, N. Transcriptional termination factors for RNA polymerase II in yeast. *Mol. Cell* **7**, 1003–1011 (2001).
59. Calvo, O. & Manley, J. L. Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination. *Mol. Cell* **7**, 1013–1023 (2001).
60. Custodio, N. *et al.* Inefficient processing impairs release of RNA from the site of transcription. *EMBO J.* **18**, 2855–2866 (1999).
61. Jensen, T. H., Patricio, K., McCarthy, T. & Rosbash, M. A block to mRNA nuclear export in *S. cerevisiae* leads to hyperadenylation of transcripts that accumulate at the site of transcription. *Mol. Cell* **7**, 887–898 (2001).
62. Gall, J. G. A role for Cajal bodies in assembly of the nuclear transcription machinery. *FEBS Lett.* **498**, 164–167 (2001).
63. Dundr, M. & Misteli, T. Functional architecture in the cell nucleus. *Biochem. J.* **356**, 297–310 (2001).
64. Sleeman, J. E. & Lamond, A. I. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Curr. Biol.* **9**, 1065–1074 (1999).
65. Misteli, T., Caceres, J. F. & Spector, D. L. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**, 523–527 (1997).
66. Misteli, T. *et al.* Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription *in vivo*. *J. Cell Biol.* **143**, 297–307 (1998).
67. Misteli, T. & Spector, D. L. RNA polymerase II targets pre-mRNA splicing factors to transcription sites *in vivo*. *Mol. Cell* **3**, 697–705 (1999).
68. Phair, R. D. & Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604–609 (2000).
69. Iborra, F. J., Pombo, A., Jackson, D. A. & Cook, P. R. Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. *J. Cell Sci.* **109**, 1427–1436 (1996).
70. Cook, P. R. The organization of replication and transcription. *Science* **284**, 1790–1795 (1999).
71. Wetterberg, I., Zhao, J., Masich, S., Wieslander, L. & Skoglund, U. *In situ* transcription and splicing in the Balbiani ring 3 gene. *EMBO J.* **20**, 2564–2574 (2001).
72. Luo, M. J. & Reed, R. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl Acad. Sci. USA* **96**, 14937–14942 (1999).
73. Le Hir, H., Moore, M. J. & Maquat, L. E. Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* **14**, 1098–1108 (2000).
74. Zhou, Z. *et al.* The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. *Nature* **407**, 401–405 (2000).
75. Le Hir, H., Izaurralde, E., Maquat, L. E. & Moore, M. J. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860–6869 (2000).
76. Luo, M. L. *et al.* Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature* **413**, 644–647 (2001).
77. Strasser, K. & Hurt, E. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yralp. *Nature* **413**, 648–652 (2001).
78. Gatfield, D. *et al.* The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. *Curr. Biol.* **11**, 1716–1721 (2001).
79. Le Hir, H., Gatfield, D., Izaurralde, E. & Moore, M. J. The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987–4997 (2001).
80. Kataoka, N. *et al.* Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol. Cell* **6**, 673–682 (2000).
81. Lykke-Andersen, J., Shu, M. D. & Steitz, J. A. Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science* **293**, 1836–1839 (2001).
82. Kim, V. N. *et al.* The Y14 protein communicates to the cytoplasm the position of exon-exon junctions. *EMBO J.* **20**, 2062–2068 (2001).
83. Kim, V. N., Kataoka, N. & Dreyfuss, G. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. *Science* **293**, 1832–1836 (2001).
84. Bachi, A. *et al.* The C-terminal domain of TAP interacts with the nuclear pore complex and promotes export of specific CTE-bearing RNA substrates. *RNA* **6**, 136–158 (2000).
85. Segref, A. *et al.* Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *EMBO J.* **16**, 3256–3271 (1997).
86. Jensen, T. H., Boulay, J., Rosbash, M. & Libri, D. The DECD box putative ATPase Sub2p is an early mRNA export factor. *Curr. Biol.* **11**, 1711–1715 (2001).
87. Reed, R. & Magni, K. A new view of mRNA export: Separating the wheat from the chaff. *Nature Cell Biol.* **3**, E201–E204 (2001).
88. Huang, Y. & Steitz, J. A. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol. Cell* **7**, 899–905 (2001).
89. Hilleren, P., McCarthy, T., Rosbash, M., Parker, R. & Jensen, T. H. Quality control of mRNA 3'-end processing is linked to the nuclear exosome. *Nature* **413**, 538–542 (2001).
90. Lei, E. P., Krebber, H. & Silver, P. A. Messenger RNAs are recruited for nuclear export during transcription. *Genes Dev.* **15**, 1771–1782 (2001).
91. Gross, S. & Moore, C. L. Rna15 interaction with the a-rich yeast polyadenylation signal is an essential step in mRNA 3'-end formation. *Mol. Cell Biol.* **21**, 8045–8055 (2001).

92. Kessler, M. M. *et al.* Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev.* **11**, 2545–2556 (1997).
93. Gonzalez, C. I., Ruiz-Echevarria, M. J., Vasudevan, S., Henry, M. F. & Peitz, S. W. The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol. Cell* **5**, 489–499 (2000).
94. Maquat, L. E. & Carmichael, G. G. Quality control of mRNA function. *Cell* **104**, 173–176 (2001).
95. Serin, G., Gersappe, A., Black, J. D., Aronoff, R. & Maquat, L. E. Identification and characterization of human orthologues to *Saccharomyces cerevisiae* Upf2 protein and Upf3 protein (*Caenorhabditis elegans* SMG-4). *Mol. Cell. Biol.* **21**, 209–223 (2001).
96. Lykke-Andersen, J., Shu, M. D. & Steitz, J. A. Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. *Cell* **103**, 1121–1131 (2000).
97. Hentze, M. W. & Kulozik, A. E. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell* **96**, 307–310 (1999).
98. Muhlemann, O. *et al.* Precursor RNAs harboring nonsense codons accumulate near the site of transcription. *Mol. Cell* **8**, 33–43 (2001).
99. Ishigaki, Y., Li, X., Serin, G. & Maquat, L. E. Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* **106**, 607–617 (2001).
100. Czaplinski, K., Ruiz-Echevarria, M. J., Gonzalez, C. I. & Peltz, S. W. Should we kill the messenger? The role of the surveillance complex in translation termination and mRNA turnover. *Bioessays* **21**, 685–696 (1999).

Acknowledgements

We thank R. Axel, D. Bentley, S. Buratowski, B. Graveley, J. Manley, M. Ptashne and members of our labs for their comments on the manuscript. We also thank R. Hellmiss for the illustrations.

Correspondence and requests for materials should be addressed to T.M. (e-mail: maniatitis@mcb.harvard.edu) or R.R. (e-mail: reed@hms.harvard.edu).

Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases

Monica Bucciantini*†, Elisa Giannoni*†, Fabrizio Chiti*§†, Fabiana Baroni*, Lucia Formigli‡, Jesús Zurdo§, Niccolò Taddei*, Giampietro Ramponi*, Christopher M. Dobson§ & Massimo Stefani*

* Dipartimento di Scienze Biochimiche, Viale Morgagni 50; ‡ Dipartimento di Anatomia, Istologia e Medicina legale, Viale Morgagni 85, Università degli Studi di Firenze, 50134 Firenze, Italy

§ Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

† These authors contributed equally to this work

A range of human degenerative conditions, including Alzheimer's disease, light-chain amyloidosis and the spongiform encephalopathies, is associated with the deposition in tissue of proteinaceous aggregates known as amyloid fibrils or plaques. It has been shown previously that fibrillar aggregates that are closely similar to those associated with clinical amyloidoses can be formed *in vitro* from proteins not connected with these diseases, including the SH3 domain from bovine phosphatidyl-inositol-3'-kinase and the amino-terminal domain of the *Escherichia coli* HypF protein. Here we show that species formed early in the aggregation of these non-disease-associated proteins can be inherently highly cytotoxic. This finding provides added evidence that avoidance of protein aggregation is crucial for the preservation of biological function and suggests common features in the origins of this family of protein deposition diseases.

An increasing body of evidence suggests that amyloid formation is the fundamental cause of protein deposition diseases^{1,2}. The nature of the pathogenic species and the mechanism by which the aggregation process results in cell damage are, however, the subject of intense debate^{3–10}. In systemic non-neurological diseases the accumulation of large quantities (sometimes kilograms) of aggregated species within a variety of organs and tissues may itself be the major cause of clinical symptoms¹¹. In other cases, particularly the degenerative neurological diseases, it appears likely that impairment of cellular function is directly linked to the interaction of protein aggregates with cellular components^{12,13}. One specific indicator of the significance of aggregate formation in pathological conditions is the evidence for a high variability in the age of disease onset¹⁴, an observation that has recently been linked to the evidence that aggregates form by nucleation¹⁵. Further clues to the molecular basis of amyloid diseases and the biological significance of protein aggregation have been provided by recent observations that a range of proteins not associated with amyloid diseases are able to aggregate *in vitro* into fibrils indistinguishable from those found in pathologic conditions^{16–19}. This finding has led to the proposal that aggregation can be viewed as a general property of polypeptide chains rather than one restricted to a small number of sequences². In the light of this conclusion we have now examined the effects on cell viability of aggregated species produced *in vitro* from two such proteins.

The SH3 domain from bovine phosphatidyl-inositol-3'-kinase (PI3-SH3) and the N-terminal ('acylphosphatase-like') domain of the *E. coli* HypF protein (HypF-N) are two examples of small globular proteins that can form fibrillar aggregates *in vitro* under appropriate conditions^{17,20}. Evidence that the aggregates formed from PI3-SH3 and HypF-N can be classified as amyloid fibrils has been obtained from electron microscopy, specific tests such as Congo red and thioflavine T binding and, in the case of PI3-SH3, X-ray diffraction^{17,20}. The heterogeneous nature of the *in vitro* aggregation process is a potential difficulty in experiments aimed at probing the nature of aggregate pathogenicity; this problem can hinder the identification of the particular species responsible for

any observed toxicity. We have found, however, that highly homogeneous populations of various types of aggregates of PI3-SH3 or HypF-N can be obtained by incubating either protein under well defined solution conditions for specific lengths of time. These sequentially and structurally unrelated proteins are therefore exceptionally favourable systems for investigating any inherent cytotoxicity of specific types of proteinaceous aggregates.

Cytotoxicity of PI3-SH3 aggregates

Incubation of PI3-SH3 in either a 50-mM acetate buffer solution, pH 5.5, containing 25% (v/v) trifluoroethanol (TFE) or a H₂O/HCl mixture, pH 2.0, results in the formation of granular or fibrillar aggregates, respectively (Fig. 1). Both types of aggregates enhance the fluorescence of thioflavin T (ThT) by factors of more than 30, indicating the presence of amyloid-like features within these structures (data not shown). Examination by transmission electron microscopy (TEM) reveals that the aggregates formed rapidly at pH 5.5 in solutions containing TFE appear as granules 4–200 nm in diameter (Fig. 1d–f) without any detectable fibrillar species. Prolonged incubation at pH 2.0, however, yields 7–13-nm-wide fibrils in the absence of any detectable granular aggregates (Fig. 1b); the width of the fibrils observed here is characteristic of the *ex vivo* amyloid fibrils extracted from patients suffering from various forms of amyloid disorders²¹. Some fibrils of this type formed from PI3-SH3 have been shown to consist of a double helical arrangement of two pairs of ribbon-like protofilaments wound around a hollow core²². In the samples analysed here, such fibrils are occasionally found to split into two 6–7-nm-wide fibrils or to associate further in a parallel fashion to produce 23- or 33-nm-wide sheet-like fibrils, also in agreement with previous findings²².

The cytotoxicity of the two types of aggregates formed in such experiments was examined by adding aliquots of the aggregates, at a range of final protein concentrations (see Fig. 1 legend), to cell culture media. Aggregate cytotoxicity was evaluated by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reduction inhibition assay, a standard indicator of cell physiological stress thought to be related to changes in intracellular trafficking,

particularly in the pathway of exocytosis^{23,24}. The experiments reveal that the highly structured PI3-SH3 fibrils formed by prolonged incubation at pH 2 do not significantly modify MTT reduction in either NIH-3T3 or PC12 cells even at the highest concentration tested (Fig. 1a). The presence of the granular aggregates formed after short incubation periods, however, significantly reduces cell viability (Fig. 1c, $P < 0.01$ at PI3-SH3 concentrations above 1 μM). No significant decrease of MTT reduction was detected when the cells were exposed either to 20 μM native PI3-SH3 or to the buffer solutions used to form the aggregates in the absence of added protein (Fig. 1).

Cytotoxicity of HypF-N aggregates

In a second series of experiments, the toxicity of aggregates formed by the protein HypF-N was examined. HypF-N was incubated at 0.3 mg ml^{-1} concentration in 50 mM acetate buffer, pH 5.5, in the presence of 30% (v/v) TFE at room temperature; aliquots of this solution were withdrawn at regular time intervals for examination by TEM and for ThT and cytotoxicity assays on cultured NIH-3T3 and PC12 cells. Under these conditions, aggregates develop within minutes; these initial aggregates exhibit both a CD (circular dichroism) spectrum indicative of β -sheet structure and an enhancement of the ThT fluorescence by over 50 times relative to that of the soluble native domain. At this stage of incubation, however, the aggregates appear completely non-fibrillar and non-granular in TEM images (Fig. 2b), whereas after 48 h of incubation some fibrillar character is clearly evident (Fig. 2c). The latter aggregates have a width of about 4–8 nm and are extremely short (typically 25–60 nm), resembling the protofibrils observed with other pro-

teins at relatively early stages of fibril formation²⁵. The ends of these aggregates appear disordered, suggesting they are in the process of undergoing a transition from amorphous to fibrillar structures. After about 20 days, aggregates of this type can no longer be detected in the samples; at this time all the aggregates appear as long unbranched fibrils with widths of either 3–5 or 7–9 nm (Fig. 2d, e), characteristic of constituent protofilaments and mature amyloid fibrils, respectively²⁰.

Addition to the cell medium of the aggregates formed after a 6 h incubation period resulted in a marked decrease in the MTT reduction by both NIH-3T3 (Fig. 2a) and PC12 cells (data not shown). This decrease is statistically highly significant at all protein concentrations tested (0.04–20 μM ; $P < 0.01$ at 0.04 μM) with respect to controls performed by incubating the same cells with 20 μM HypF-N in its soluble form or with the buffer solutions in the absence of protein. A series of experiments revealed that aggregate toxicity depends on protein concentration and initially increases with the length of exposure to the conditions that result in aggregation, reaching a maximum after 48 h (Fig. 2a). Incubating cells with protein aggregates formed at this time results in an inhibition of MTT reduction ranging from 20% at the lowest protein concentration used here (0.04 μM) to 70% at the highest protein concentration (20 μM) with respect to the control experiments. In addition to impairing cellular function, HypF-N aggregates were also found to lead to cell death, as revealed by the trypan blue internalization test (Fig. 3). The results show that the rate of cell mortality increases until it reaches nearly 40% at the highest protein concentrations, indicating that the cellular dysfunction shown by the MTT assay leads to cell death. Addition to the cell medium of

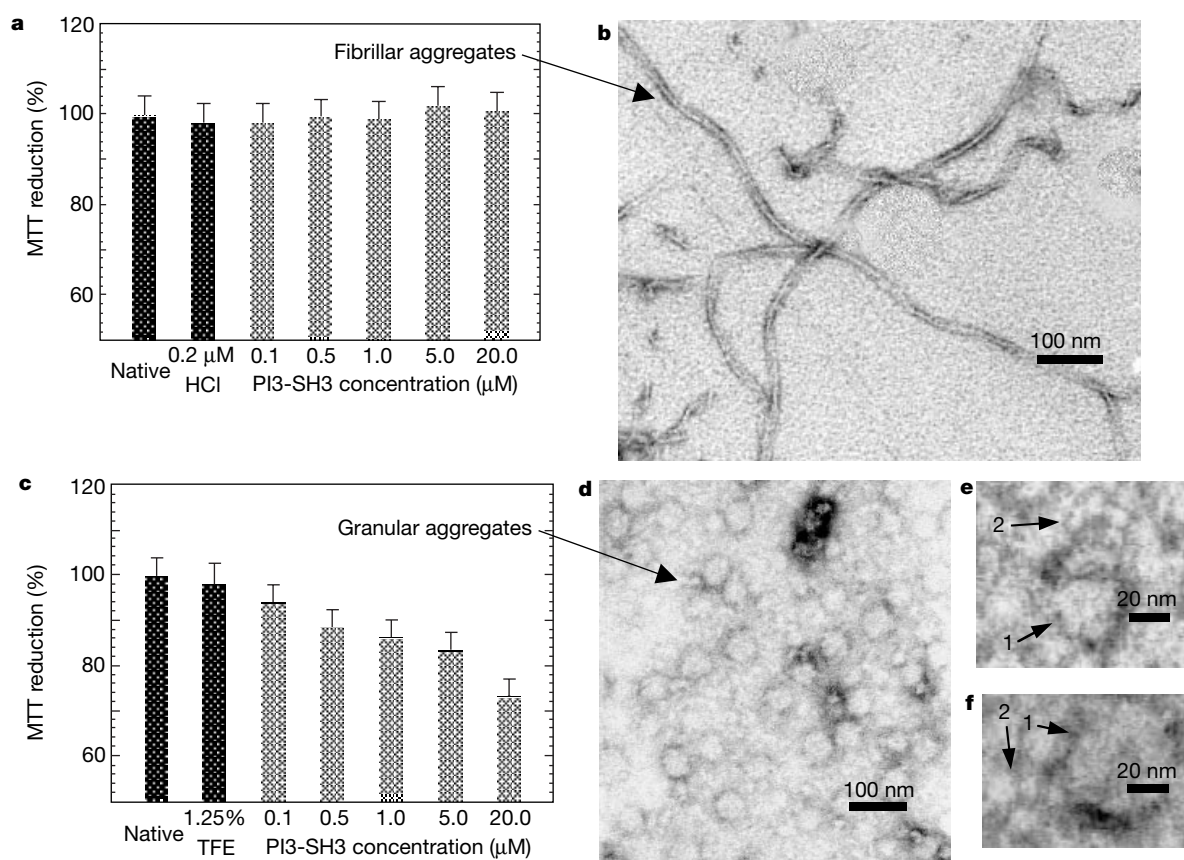


Figure 1 PI3-SH3 aggregation and cytotoxicity. **a**, **c**, Cytotoxic effect of fibrillar (**a**) and granular (**c**) PI3-SH3 aggregates on NIH-3T3 cells. Reported values are those after treatment with 20 μM native protein (Native), with cell medium containing an aliquot of the solution where aggregates were formed (HCl or TFE) or with aggregates at the indicated

protein concentrations (PI3-SH3). Values are relative to those of control cells treated with complete medium alone. **b**, **d–f**, Electron micrographs showing fibrillar (**b**) and granular (**d–f**) aggregates formed from PI3-SH3. Labels 1 and 2 indicate large granules and clusters of small granules, respectively (**e**, **f**).

protein samples incubated for times longer than 48 h, however, resulted in a progressive decrease of cell impairment that correlates, at least qualitatively, with the increase in the proportion of mature fibrils relative to other aggregates in the protein samples. Even at the highest protein concentrations analysed, MTT reduction approaches the value of the controls when the cells were exposed to protein samples incubated for 21 days (Fig. 2a), conditions where only mature fibrils are evident in the TEM micrographs. A lack of toxicity was also found when cells were exposed to well defined amyloid fibrils formed at low pH (50 mM citric acid, pH 3.0) (data not shown).

The origins of aggregate toxicity

The results of the present study indicate that aggregates formed from PI-SH3 and HypF-N can be substantially cytotoxic. In both cases, the cytotoxicity was found to depend on the supramolecular organization of the amyloid aggregates and is much more pronounced for the rapidly formed non-fibrillar aggregates than for the highly organized fibrillar structures. The results for these two proteins, neither of which is disease-related, are very similar to those reported for the disease-associated fusogenic prion protein fragment, for α -synuclein, for A β (1–42) and for transthyretin^{3–10,26}. Remarkably, the levels of cell impairment induced by the most toxic species found in the present study are comparable to those of the highly toxic aggregates formed from A β (1–42) (ref. 27). The data, therefore, suggest that impairment of cell viability by protein aggregates of the type that can subsequently form amyloid fibrils could be a general phenomenon and not simply a specific property of the small number of polypeptides associated with clinically recognized protein deposition diseases. This result is of particular significance in the light of the recent conclusion that the ability to form highly

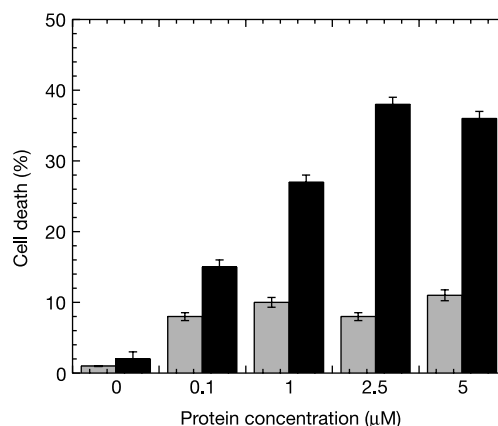


Figure 3 Percentage of cell deaths induced by 48-h-aged HypF-N aggregates at different protein concentrations. Solid bars refer to aggregated protein, dotted bars to control experiments performed in the presence of soluble protein (see Methods for details).

ordered amyloid fibrils is itself a generic property of proteins^{17–19}. In addition, the mature fibrils of the proteins we tested appear to be essentially harmless to cells, providing an explanation for previous data indicating that well defined amyloid fibrils formed by a synthetic peptide are not cytotoxic²⁷ and reinforcing the findings on the cytotoxicity of A β , α -synuclein and transthyretin pre-fibrillar assemblies both in cultured cells and in transgenic mice^{3–5,9,10,26,10,28}.

The results reported in this paper, therefore, provide evidence that protein aggregates not associated with disease can be inherently cytotoxic and result in substantial impairment of cellular function

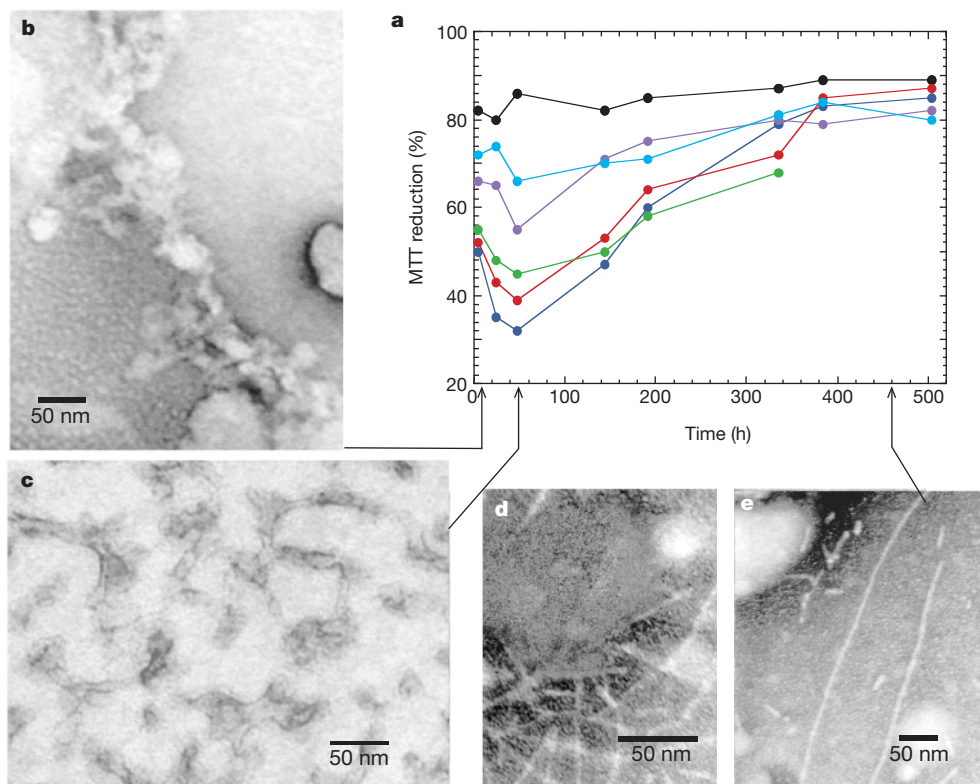


Figure 2 HypF-N aggregation and cytotoxicity. **a**, Differential toxicity of amorphous, protofibrillar and fibrillar HypF-N aggregates. Cell viability was checked by the MTT inhibition reduction test, after addition to the cell medium of either 20 μM native protein (black circles) or different concentrations of protein incubated in TFE: 20 μM (blue), 5 μM (red), 1 μM (green), 0.2 μM (purple) and 0.04 μM (cyan). Values are relative to control

cells treated with complete medium alone. The electron micrographs show amorphous aggregates of HypF-N, formed after 6 h incubation (**b**), amorphous aggregates developing into fibrils after 48 h incubation (**c**) and mature amyloid protofilaments (**d**) and fibrils (**e**) after 20 days incubation.

or even cell death. There is a considerable debate as to whether fully formed mature amyloid fibrils or rapidly formed aggregates that precede their formation are the primary pathogenic species responsible for the onset of disease^{3–10}. Although there is evidence for the toxicity of mature fibrils in some amyloid diseases, the data reported here support previous suggestions that, at least in some cases, the non-fibrillar aggregates that precede formation of mature amyloid fibrils may be the primary toxic species^{3–5,9,10,26,10,28}. This toxicity is likely to arise because in these early aggregates hydrophobic side-chains and other regions of the polypeptide chain will be much more accessible than in the fully formed mature fibrils. Indeed, the latter are often found to be remarkably inert, for example in their resistance to proteolysis and degradation²⁹.

The inherent cytotoxicity of the aggregates formed by the two non-disease-related proteins studied here suggests not only that the toxicity of aggregated species could be a general phenomenon, but also that the pathogenicity of protein-deposition diseases could be primarily related to the structural nature of the aggregates rather than to the specific sequences of the proteins from which they arise. We suggest that toxicity could primarily arise because on the surface of disordered aggregates there is likely to be a combinatorial display of amino acids enabling these species to interact inappropriately with a wide range of cellular components. Such a conclusion further suggests that the differing clinical manifestations of amyloid formation, ranging from neuronal cell death to the accumulation of large quantities of proteinaceous material, could arise in large part because of variations in the nature and morphologies of the specific aggregates as they form in the different diseases, as well as on the susceptibility of different cell types to the various aggregates. Such variations will occur as a consequence of the specific character of the different proteins involved and their location, and also on the different conditions under which aggregation takes place.

Implications for misfolding diseases and biological evolution

The present findings, that early aggregates formed by a wider range of proteins than those known to be associated with neurological diseases can be cytotoxic, provide new opportunities to define the nature of amyloid diseases and the mechanism of amyloid toxicity at the molecular level. They also raise the possibility that trace amounts of aggregates of a variety of proteins might occur spontaneously, particularly during ageing, and that such aggregates could account for subtle impairments of cellular function in the absence of an evident amyloid phenotype. It would thus be interesting to search for early protein aggregates in systemic and neurologic disorders not presently associated with amyloid formation. More generally, knowledge of the origin and nature of aggregate pathogenicity is of crucial importance in efforts to identify the correct targets for drug design in the search for effective therapeutic protocols.

The inherent toxicity in protein aggregates could also help us to understand fundamental aspects of cell biology. It suggests, for example, that avoidance of aggregation could be more important for the proper functioning of biological organisms than was previously suspected if aggregates of proteins are often toxic rather than simply non-functional. In this case, in addition to increasing the efficiency of folding and rescuing misfolded proteins after biosynthesis³⁰, evolutionary developments to prevent aggregate formation, notably molecular chaperones, ubiquitination enzymes and proteasomes, are needed for the preservation of the long-term viability of living organisms. This latter idea is reinforced by recent findings concerning the relationships between neurodegenerative diseases and failure of cellular defence mechanisms targeted towards misfolded proteins and the existence, in both prokaryotic and eukaryotic cells, of a complex regulatory system of intracellular protein degradation (see ref. 31 and references therein). Such results, together with findings that aggregate formation is linked to the inheritance of specific traits in organisms such as yeasts³²,

provides increasing evidence that the control of protein misfolding and aggregation in addition to being of fundamental importance for cell viability, has been a major driving force in biological evolution. □

Methods

Protein aggregate production

PI3-SH3 was expressed and purified as previously reported³³. Granular aggregates were formed incubating the protein for 1 h at 20 °C at a concentration of 10 mg ml⁻¹ in 50 mM acetate buffer, pH 5.5, containing 25% (v/v) TFE. Fibrillar aggregates were grown for 1 month at a protein concentration of 10 mg ml⁻¹ in a water/HCl mixture, pH 2.0, at 37 °C. Conditions for HypF-N purification and aggregation have been described previously²⁰.

Cell culture

NIH-3T3 cells (mouse fibroblasts, American Type Culture Collection) were routinely cultured in Dulbecco's modified Eagle's medium (Gibco BRL) containing 10.0% bovine calf serum and 3.0 mM glutamine in a 5.0% CO₂ humidified environment, at 37 °C. PC12 cells (rat pheochromocytoma, American Type Culture Collection) were cultured in RPMI medium (Gibco BRL) supplemented with 10.0% horse serum, 5.0% fetal bovine serum and 3.0 mM glutamine in a 5.0% CO₂ humidified atmosphere at 37 °C. 100 U ml⁻¹ penicillin and 100 µg ml⁻¹ streptomycin were added to both media. Cells were used for a maximum of 20 passages. NIH-3T3 or PC12 cells were plated at a density of 10,000 cells per well on 96-well plates in 100 µl of fresh medium. After 24 h, the NIH-3T3 medium was exchanged with 100 µl of DMEM, without phenol red, containing 10.0% bovine calf serum, and the PC12 medium was changed with 100 µl of RPMI, without phenol red, supplemented with 10.0% horse serum and 5.0% fetal bovine serum.

Incubation of cells in the presence of protein aggregates

Aliquots of solutions containing native or aggregated proteins were centrifuged, dried under N₂ to remove TFE when necessary, dissolved in RPMI without phenol red and immediately added to the cell media at 0.1–20 µM (PI3-SH3) or 0.04–20.0 µM (HypF-N) final concentrations. After 24 h incubation, 10 µl of a stock MTT solution in PBS was added to give a final concentration of 0.5 mg ml⁻¹ and incubated for a further 4 h. 100 µl of cell lysis buffer (20.0% SDS, 50.0% N,N-dimethylformamide, pH 4.7) was added to each well and the samples were incubated overnight at 37 °C in an humidified incubator. Absorbance values of blue formazan were determined at 590 nm with an automatic plate reader. Cell death was assessed by the trypan blue internalization test³⁴. After a 24 h incubation with 48 h aged HypF-N in TFE or with the soluble domain, NIH-3T3 cells were treated with trypan blue and survival was quantified by counting (three fields per well, two wells per condition, an average of 50 cells per field).

Electron microscopy

TEM images were acquired using a JEM 1010 transmission electron microscope at 80 kV excitation voltage. In each case, 3.0 µl of protein solution was placed on a formvar and carbon-coated grid and blotted off after 3 min. The sample was then stained with 3 µl of 2.0% uranyl acetate, dried and observed at a magnification of 12–30,000.

ThT staining

60 µl aliquots of protein solution were mixed with 0.44 ml of 25 µM ThT in 25 mM phosphate buffer, pH 6.0, and the resulting fluorescence measured immediately after mixing using a Shimadzu RF-5000 spectrofluorimeter at excitation and emission wavelengths of 440 and 485 nm, respectively.

Received 2 January; accepted 11 February 2002.

- Kelly, J. W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **8**, 101–106 (1998).
- Dobson, C. M. The structural basis of protein folding and its links with human disease. *Phil. Trans. R. Soc. Lond. B* **356**, 133–145 (2001).
- Lambert, M. P. et al. Diffusible, nonfibrillar ligands derived from Aβ₄₂ are potent central nervous system neurotoxins. *Proc. Natl Acad. Sci. USA* **95**, 6448–6453 (1998).
- Hartley, D. M. et al. Protofibrillar intermediates of amyloid beta-protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. *J. Neurosci.* **19**, 8876–8884 (1999).
- Pillot, T. et al. The nonfibrillar amyloid β-peptide induces apoptotic neuronal cell death: involvement of its C-terminal fusogenic domain. *J. Neurochem.* **73**, 1626–1634 (1999).
- Monji, A. et al. Inhibition of Aβ fibril formation and Aβ-induced cytotoxicity by senile plaque-associated proteins. *Neurosci. Lett.* **278**, 81–84 (2000).
- Walsh, D. M. et al. Amyloid β-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates. *J. Biol. Chem.* **274**, 25945–25952 (1999).
- Goldberg, M. S. & Lansbury, P. T. Jr Is there a cause-and-effect relationship between alpha-synuclein fibrillization and Parkinson's disease? *Nature Cell Biol.* **2**, E115–E119 (2000).
- Conway, K. A. et al. Acceleration of oligomerization, not fibrillization, is a shared property of both alpha-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proc. Natl Acad. Sci. USA* **97**, 571–576 (2000).
- Zhu, Y. J., Lin, H. & Lal, R. Fresh and nonfibrillar amyloid beta protein (1–40) induces rapid cellular degeneration in aged human fibroblasts: evidence for A beta P-channel-mediated cellular toxicity. *FASEB J.* **14**, 1244–1254 (2000).
- Pepys, M. B. in *Oxford Textbook of Medicine* (eds Weatherall, D. J., Ledingham, J. G. & Warrel, D. A.) 3rd edn, 1512–1524 (Oxford Univ. Press, Oxford, 1995).
- Lorenzo, A. & Yankner, B. A. β-amyloid neurotoxicity requires fibril formation and is inhibited by congo red. *Proc. Natl Acad. Sci. USA* **91**, 12243–12247 (1994).

13. Thomas, T., Thomas, G., McLendon, C., Sutton, T. & Mullan, M. β -Amyloid-mediated vasoactivity and vascular endothelial damage. *Nature* **380**, 168–171 (1996).
14. Clarke, G. *et al.* A one-hit model of cell death in inherited neuronal degenerations. *Nature* **406**, 195–199 (2000).
15. Perutz, M. F. & Windle, A. H. Cause of neuronal death in neurodegenerative disease attributable to expansion of glutamine repeats. *Nature* **412**, 143–144 (2001).
16. Glenner, G. G., Eanes, E. D., Bladen, H. A., Linke, R. P. & Termine, J. D. Beta-pleated sheet fibrils. A comparison of native amyloid with synthetic protein fibrils. *J. Histochem. Cytochem.* **22**, 1141–1158 (1974).
17. Guijarro, J. I., Sunde, M., Jones, J. A., Campbell, I. D. & Dobson, C. M. Amyloid fibril formation by an SH3 domain. *Proc. Natl Acad. Sci. USA* **95**, 4224–4228 (1998).
18. Chiti, F. *et al.* Designing conditions for *in vitro* formation of amyloid protofilaments and fibrils. *Proc. Natl Acad. Sci. USA* **96**, 3590–3594 (1999).
19. Fandrich, M., Fletcher, M. A. & Dobson, C. M. Amyloid fibrils from muscle myoglobin. *Nature* **410**, 165–166 (2001).
20. Chiti, F. *et al.* Solution conditions can promote formation of either amyloid protofilaments or mature fibrils from the HypF N-terminal domain. *Protein Sci.* **10**, 2541–2547 (2001).
21. Sunde, M. & Blake, C. F. The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Protein Chem.* **50**, 123–159 (1997).
22. Jimenez, J. L. *et al.* Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing. *EMBO J.* **18**, 815–821 (1999).
23. Liu, Y., Peterson, D. A., Kimura, H. & Schubert, D. Mechanism of cellular 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reduction. *J. Neurochem.* **69**, 581–593 (1997).
24. Abe, K. & Saito, H. Amyloid beta protein inhibits cellular MTT reduction not by suppression of mitochondrial succinate dehydrogenase but by acceleration of MTT formazan exocytosis in cultured rat cortical astrocytes. *Neurosci. Res.* **31**, 295–305 (1998).
25. Harper, J. D., Lieber, C. M. & Lansbury, P. T. Jr Atomic force microscopy imaging of seeded fibril formation and fibril branching by Alzheimer's disease amyloid- β protein. *Chem. Biol.* **4**, 951–959 (1997).
26. Mendes Sousa, M., Cardoso, I., Fernandes, R., Guimaraes, A. & Saraiva, M. J. Deposition of transthyretin in early stages of familial amyloidotic polyneuropathy. Evidence for toxicity of nonfibrillar aggregates. *Am. J. Pathol.* **159**, 1993–2000 (2001).
27. Fezoui, Y. *et al.* A de novo designed helix-turn-helix peptide forms non-toxic amyloid fibrils. *Nature Struct. Biol.* **7**, 1095–1099 (2000).
28. Hsia, A. Y. *et al.* Plaque-independent disruption of neural circuits in Alzheimer's disease mouse models. *Proc. Natl Acad. Sci. USA* **96**, 3228–3233 (1999).
29. Zurdo, J., Guijarro, J. I. & Dobson, C. M. Preparation and characterisation of purified amyloid fibrils. *J. Am. Chem. Soc.* **123**, 8141–8142 (2001).
30. Leroux, M. R. & Hartl, F. U. in *Mechanisms of Protein Folding* (ed. Pain, R. H.) 2nd edn, 364–405 (Oxford Univ. Press, Oxford, 1999).
31. Sherman, M. Y. & Goldberg, A. L. Cellular defenses against unfolded proteins: a cell biologist thinks about neurodegenerative diseases. *Neuron* **29**, 15–32 (2001).
32. Li, L. R. & Lindquist, S. Creating a protein-based element of inheritance. *Science* **287**, 661–664 (2000).
33. Booker, J. W. *et al.* Solution structure and ligand-binding site of the SH3 domain of the P85 α -subunit of phosphatidylinositol-3-kinase. *Cell* **73**, 813–822 (1993).
34. Butterfield, D. A., Yatin, S. M., Varadarajan, S. & Koppal, T. Amyloid β -peptide-associated free radical oxidative stress, neurotoxicity and Alzheimer's disease. *Methods Enzymol.* **309**, 746–768 (1999).

Acknowledgements

The work was supported by grants from the Italian MIUR (PRIN “Folding e Misfolding di Proteine”) and the Italian Telethon Foundation. The research of C.M.D. is supported in part by a Programme Grant from the Wellcome Trust. F.C. is supported by a fellowship from the Italian Telethon Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.S. (e-mail: stefani@scibio.unifi.it) or C.M.D. (e-mail: cmd44@cam.ac.uk).

The signature of supernova ejecta in the X-ray afterglow of the γ -ray burst 011211

J. N. Reeves*, D. Watson*, J. P. Osborne*, K. A. Pounds*, P. T. O'Brien*
A. D. T. Short*, M. J. L. Turner*, M. G. Watson*, K. O. Mason†, M. Ehle‡
& N. Schartel‡

*X-ray Astronomy Group, Department of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK

†Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking RH5 6NT, UK

‡XMM-Newton SOC, Villafranca, 28080, Madrid, Spain

Now that γ -ray bursts (GRBs) have been determined to lie at cosmological distances¹, their isotropic burst energies are estimated to be as high as 10^{54} erg (ref. 2), making them the most energetic phenomena in the Universe. The nature of the progenitors responsible for the bursts remains, however, elusive. The favoured models range from the merger of two neutron stars in a binary system^{3–5} to the collapse of a massive star^{6–8}. Spectroscopic studies of the afterglow emission could reveal details of the environment of the burst, by indicating the elements present, the speed of the outflow and an estimate of the temperature. Here we report an X-ray spectrum of the afterglow of GRB011211, which shows emission lines of magnesium, silicon, sulphur, argon, calcium and possibly nickel, arising in metal-enriched material

with an outflow velocity of the order of one-tenth the speed of light. These observations strongly favour models³⁰ where a supernova explosion from a massive stellar progenitor precedes the burst event and is responsible for the outflowing matter.

The γ -ray burst GRB011211 was first detected on 11 December 2001 at 19:09:21 UT, by the Beppo-SAX satellite⁹; the burst duration was 270 s (making GRB011211 the longest burst observed by Beppo-SAX), with a peak flux (40–700 keV) of 5×10^{-8} erg cm⁻² s⁻¹. Spectroscopy of the optical afterglow revealed several absorption lines at a redshift of $z = 2.141 \pm 0.001$ (refs 10, 11), and R-band imaging¹² has linked the optical transient with extended emission—the probable host galaxy—of magnitude $m_v = 25.0 \pm 0.5$. Assuming the absorption system arises from the GRB host galaxy, and adopting a cosmology of $H_0 = 75$ km s⁻¹ Mpc⁻¹ and $q_0 = 0.1$, implies a total equivalent isotropic energy for GRB011211 of 5×10^{52} erg.

The observations of GRB011211 by the orbiting XMM-Newton¹³ X-ray telescope started at 06:16:56 UT on 12 December 2001, 11 hours after the initial burst¹⁴. Data from the European Photon Imaging Cameras (EPIC) have been analysed, using both the MOS and PN instruments; the total observation duration is 27 ks. The time-averaged 0.2–10 keV flux, F , was 1.9×10^{-13} erg cm⁻² s⁻¹ (corresponding to a 0.6–30 keV rest-frame afterglow luminosity of 7×10^{45} erg s⁻¹), decreasing with time t during the observation as $F(t) \propto t^{-(1.7 \pm 0.2)}$. The Optical Monitor detected the burst afterglow in both the visible and ultraviolet (UVW1) bands (at 11 σ and 6 σ confidence), with magnitudes of 21.12 ± 0.13 and 21.6 ± 0.3 , respectively. The magnitudes correspond to fluxes of $(1.28 \pm 0.16) \times 10^{-17}$ erg cm⁻² s⁻¹ Å⁻¹ and $(0.85 \pm 0.23) \times 10^{-17}$ erg cm⁻² s⁻¹ Å⁻¹ at wavelengths of 5,430 Å and 3,140 Å,

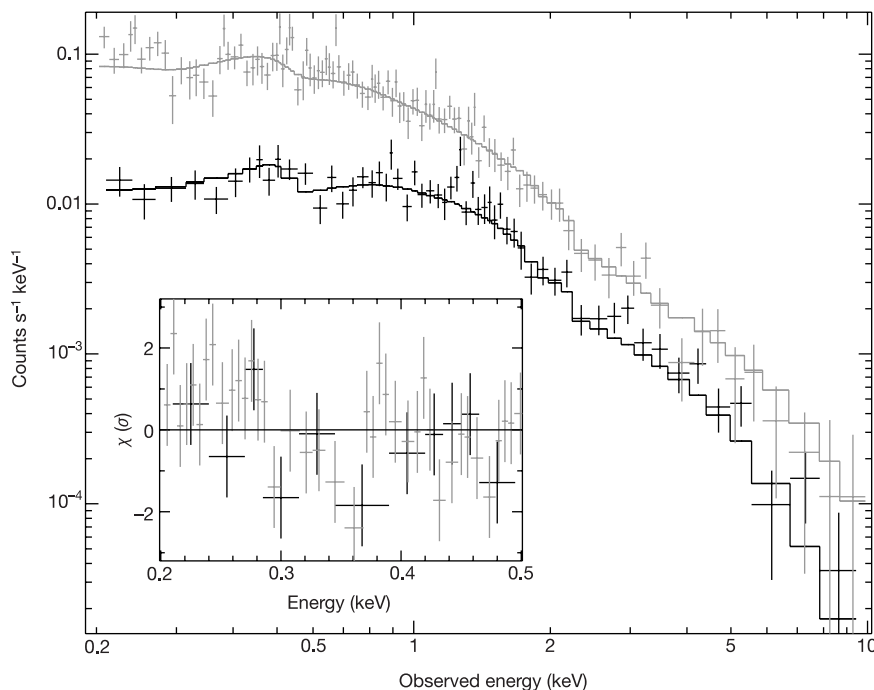


Figure 1 The XMM-Newton spectrum of the afterglow of the γ -ray burst, GRB011211. The spectrum shown is integrated over the whole 27-ks observation. EPIC-PN detector data points are shown in grey, EPIC-MOS detector points in black (see text). The ordinate shows the counts per second obtained in each detector, over a given energy range (keV), and the abscissa plots the observed energy (in the observed frame) of the X-ray photons. The solid curves show the model fit, convolved through the instrument response and energy resolution. In this case, the continuum emission can be modelled by either a power law of photon index $\Gamma = 2.3 \pm 0.1$ or thermal bremsstrahlung emission of temperature $kT = 2.1 \pm 0.2$ keV, attenuated only by an absorption column of 4.2×10^{20} cm⁻² from

our own Galaxy. An apparent absorption feature is present between 0.3 and 0.4 keV (see inset); this can be modelled by an absorption edge (where the opacity after the edge energy E_0 falls as E^{-3}). In the burst rest frame, the edge energy is $E_0 = 0.99 \pm 0.09$ keV with an optical depth $\tau = 1.0 \pm 0.4$. Assuming the above absorption feature arises in the same material as the line emission, then its likely identification is with the O VIII edge at 0.87 keV. Note however that although the absorption edge is formally significant (99.9% from an F -test), the detector resolution is lowest at this energy ($\Delta E \approx 100$ eV at 300 eV), which makes the shape of the feature difficult to constrain.

respectively. These data are consistent with reddening in the host galaxy of $E(B - V) \approx 0.2$, or with Lyman α absorption at high redshift. Figure 1 shows the time-averaged EPIC X-ray spectrum of the burst afterglow. The continuum can be fitted by a power law of photon index $\Gamma = 2.3 \pm 0.1$, attenuated by the Galactic absorption column of $4.2 \times 10^{20} \text{ cm}^{-2}$. An apparent absorption feature is present between 0.3 and 0.4 keV, which, when fitted by an absorption edge, gives an energy of $E_0 = 0.99 \pm 0.09 \text{ keV}$ and an optical depth of $\tau = 1.0 \pm 0.4$, in the burst rest frame.

Spectra were initially extracted in five time segments, corresponding to 0–5 ks, 5–10 ks, 10–15 ks, 15–20 ks and 20–27 ks after the start of the XMM-Newton observation. As significant line emission was observed only during the initial 10 ks, the remaining data were binned into one 17-ks time bin. Figure 2 shows the EPIC-PN spectrum between 0.2 and 3 keV, from the first 5 ks only; emission lines are detected in the burst rest frame at energies (keV) of 1.40 ± 0.05 , 2.19 ± 0.04 , 2.81 ± 0.04 , 3.79 ± 0.07 and 4.51 ± 0.12 . The closest abundant K α transitions to the observed lines are: Mg XI (1.35 keV) or Mg XII (1.47 keV), Si XIV (2.00 keV), S XVI (2.62 keV), Ar XVIII (3.32 keV) and Ca XX (4.10 keV). Thus we infer that the lines are blueshifted with respect to the known redshift of GRB011211. To quantify this, we adopted the rest-frame energies corresponding to Mg XI, Si XIV, S XVI, Ar XVIII and Ca XX, and varied the redshift of the line-emitting material. The best-fit redshift for the line set was found to be $z = 1.88 \pm 0.06$, differing significantly (at $>99.99\%$ confidence) from the known GRB redshift of $z = 2.140 \pm 0.001$, implying an outflow velocity (v) for the line emitting material of $v/c = 0.086 \pm 0.004$, where c is the velocity of light, or $v = 25,800 \pm 1,200 \text{ km s}^{-1}$. The line emission also declines

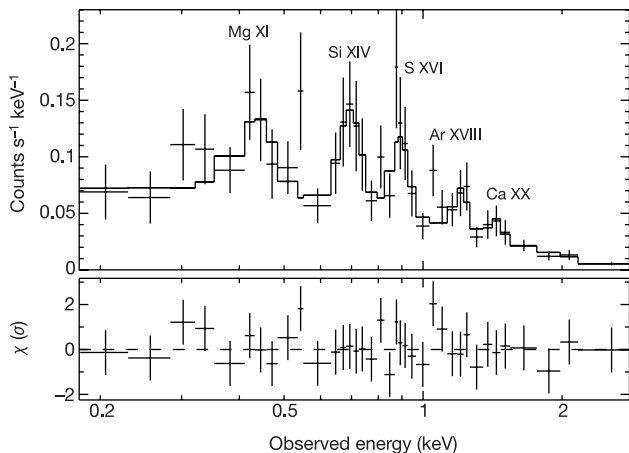


Figure 2 The XMM-Newton EPIC-PN spectrum of the burst afterglow, for the first 5 ks of exposure only. Top, the observed count rate spectrum; bottom, the residuals of the thermal model compared with the data points, in units of 1σ deviations. The energy plotted on the abscissa is in the observer frame. Emission lines are detected at energies (keV) of 0.45 ± 0.03 , 0.70 ± 0.02 , 0.89 ± 0.01 , 1.21 ± 0.02 and 1.44 ± 0.04 in the observed spectrum, whilst the measured line fluxes ($\text{erg cm}^{-2} \text{ s}^{-1}$) are $(7.6 \pm 5.1) \times 10^{-15}$, $(1.1 \pm 0.3) \times 10^{-14}$, $(9.9 \pm 2.9) \times 10^{-15}$, $(6.7 \pm 2.5) \times 10^{-15}$ and $(4.4 \pm 2.2) \times 10^{-15}$, respectively. These correspond to energies (keV) of 1.40 ± 0.05 , 2.19 ± 0.04 , 2.81 ± 0.04 , 3.79 ± 0.07 and 4.51 ± 0.12 in the burst rest frame (at $z = 2.14$), with rest-frame equivalent widths (eV) of 180 ± 120 , 430 ± 130 , 480 ± 140 , 460 ± 170 and 360 ± 180 , respectively (all errors are quoted at 1σ confidence). Note that an iron K line is not detected; the upper limit on the rest-frame equivalent width is $<400 \text{ eV}$. For reference, the resolution (full-width at half-maximum) of the EPIC-PN spectrum is 100 eV at 1 keV . The emission lines can be identified with the K α transitions of Mg XI (or Mg XII), Si XIV, S XVI, Ar XVIII, and Ca XX. The observed energies are blue-shifted by a factor corresponding to a velocity of $v = 0.086c$ (or $25,800 \text{ km s}^{-1}$), when compared to the redshift of the γ -ray burst (at $z = 2.14$). For clarity, only the EPIC-PN data are shown; consistent results are obtained for the EPIC-MOS camera, although the signal-to noise ratio of the MOS data is lower.

more rapidly than the continuum (at $>3\sigma$ confidence), suggesting a more enduring (non-thermal) component to the continuum flux. The decrease in flux of the Si XIV line is shown in Fig. 3; the line is detected only over the first 10 ks of observation.

In order to assess the quality of the spectral fit and the statistical significance of the emission lines, we calculated the fit statistic, measured in terms of total χ^2 deviations between the data points and the input model, divided by the degrees of freedom (d.o.f) in the fit. This improved from $\chi^2/\text{d.o.f} = 56.7/47$ for a pure power-law model to $36.2/41$ on the addition of the lines. Employing an F -test¹⁵ then yields a significance level of 99.7% for the set of lines as a whole. Furthermore, by performing Monte Carlo simulations, we find a probability of only 0.02% that the lines result from purely random Poisson noise. We conclude that the line emission is detected with good confidence.

X-ray line emission can arise by a variety of processes, including thermal emission, recombination in a photoionized plasma, or by reflection of hard X-rays in dense, optically thick matter. Modelling the early spectrum with emission from an optically thin plasma of temperature $kT = 4.5 \pm 0.5 \text{ keV}$ and luminosity $7 \times 10^{45} \text{ erg s}^{-1}$ (using the 'MEKAL' code¹⁶), requires an over-abundance for Mg, Si, S, Ar and Ca of about 10 times the solar value (solar abundance is ruled out at $>99.9\%$ confidence). The absence of emission in the Fe K band gives a limit to the abundance of iron at <1.3 times the solar value. The fit statistic for the thermal model is excellent ($\chi^2/\text{d.o.f} = 37/44$), and Monte Carlo simulations showed that only one in 10,000 pure power-law spectra could yield as good a fit by chance alone (that is, a null probability of about 0.01%). Another possible source of the line emission is X-ray reflection^{17,18}. This is currently favoured in 'nearby reprocessor' models¹⁹, where X-rays scatter off dense material within the stellar envelope of a massive progenitor star. We fitted ionized reflection models¹⁷ to the data for two cases: (1) where the walls of the stellar envelope subtend $2\pi \text{ sr}$ solid angle to the X-ray source; (2) where the emission arises purely from the scattered X-ray flux, with no continuum emission. In both cases the fit obtained is poor ($\chi^2/\text{d.o.f} = 62.4/46$ and $56.7/47$, respectively);

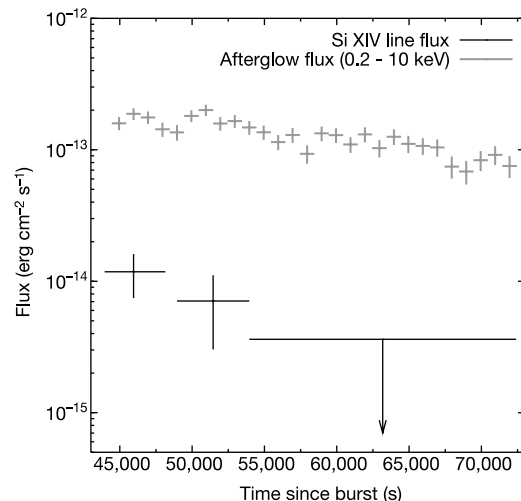


Figure 3 The line flux of Si XIV K α , and the total continuum flux (0.2–10 keV), as a function of time since the initial burst. The Si XIV line is detected in the first 10 ks of the observation (illustrated by the first two data points), but is not detected during the remainder of the observation (the upper limit shown is at the 3σ confidence level). The values of the Si XIV line flux are $(1.1 \pm 0.3) \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ during the first 5 ks, $(0.7 \pm 0.4) \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ from 5 to 10 ks and $<3.6 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ during the last 17 ks of observation. A similar effect is seen for the other emission lines, none being detected in the later part of the observation. For instance, the S XVI line flux is $(1.0 \pm 0.3) \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the first 5 ks, decreasing to $<2.5 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$ during the last 17 ks of observation.

the reflection models fail to reproduce the emission lines observed in the early afterglow spectrum.

A major drawback of both the photoionization and reflection models is that they favour emission from heavier metals such as Fe or Ni, rather than the lighter elements (because the line fluxes increase as the square of the atomic number), which is difficult to reconcile with the observed line spectrum. Therefore we consider the thermal model to be physically plausible, noting that the non-detection of Fe line emission is then a less severe constraint. We assume that a dense shell of material (the ejecta of a very recent supernova) is heated by the γ -ray burst. At the densities deduced below, the emitting plasma will rapidly reach thermal equilibrium, and will radiate over a timescale given by $t_{\text{cool}} \approx 2 \times 10^{15} n_e^{-1} T_8^{1/2}$ s, where T_8 is the plasma temperature in units of 10^8 K, and n_e is the electron density (in cm^{-3}). From the observed X-ray luminosity, we estimate an emission measure $n_e^2 V = 10^{69} \text{ cm}^{-3}$, where V is the total emitting volume, whilst the temperature is also derived from the thermal spectrum (4.5 keV or 5×10^7 K). We interpret the observed emission line timescale (t) in terms of the light travel delay in radiation arising from different parts of the illuminated shell. For a spherical shell this delay depends on the radius of the shell R and the beam half-opening angle θ , and is given by $R = ct/(1 - \cos \theta)$. Converting the duration of the line emission to the burst rest-frame, $t = 10^4/(1+z)$ s, for a half-opening angle of $\theta \approx 20^\circ$ (consistent with GRB beaming models²⁰), yields $R \approx 10^{15}$ cm. Combining the emission measure of illuminated matter, with an electron scattering optical depth of approximately 1 for the shell, then allows estimates (for $R \approx 10^{15}$ cm) of the electron density $n_e \approx 10^{15} \text{ cm}^{-3}$ and shell thickness $\sim 10^9$ cm (in practice the ejecta may be clumpy, and thus distributed over a thicker shell). The mass of the illuminated matter is then approximately 0.1 solar masses ($\sim 10^{32}$ g) with a kinetic energy of $\sim 10^{51}$ erg, consistent with the energies of a typical supernova and the γ -ray burst (if beamed). In the isotropic limit, where the thermal emission results from a spherical shell of solid angle 4π sr, then the radius of the shell will simply be $R = ct$, yielding $R \approx 10^{14}$ cm; this sets a minimum distance between the burst and the line-emitting material. We might also expect to see redshifted line emission from the counter-jet receding from us, but matter outside the beam may prevent the X-ray emission from the far side of the shell being seen.

With a mean outflow velocity of the ejecta of 0.1c, the estimated radius of the shell, $R \sim 10^{15}$ cm, implies a time delay between the supernova and γ -ray burst of about 4 days (in the isotropic limit, this time delay will be a minimum of 10 hours). This offers a natural explanation for the non-detection of Fe K emission in the XMM-Newton spectrum. The stable isotope ^{56}Fe is produced by β -decay via the reaction $^{56}\text{Ni} \rightarrow ^{56}\text{Co} \rightarrow ^{56}\text{Fe}$, the reactions having half-lives of 6 and 78 days, respectively. Over a short timescale after the supernova, as our model implies, very little Fe will have been synthesized. Instead, we would expect to observe spectral features due to Ni or Co. We note that we do find a marginal detection (90% confidence) of the blueshifted Ni K α emission line, with a rest-frame equivalent width of 800 eV (equivalent to an eventual iron abundance of four times the solar value). Although iron will be present in the collapsed core of the massive progenitor, most of the illuminated ejecta will be from the outer stellar layers (that is, the lower-atomic-number elements). Our result contrasts with the large masses of iron implied by the iron lines reported in the X-ray spectra of several other afterglows^{21–23}. It seems plausible that in some of those cases, the actual detection was of Ni K α ²⁴; otherwise a much longer delay between an initial supernova and the onset of the γ -ray burst is required, of the order of several months.

The XMM-Newton observations reported here seem to rule out a neutron-star binary merger as the progenitor of GRB011211, as such a merger is unlikely to expel sufficient quantities of matter into the surrounding medium²⁵, nor can the relatively low iron abun-

dance be explained, as any supernova will have occurred many years before the stellar merger. The X-ray-emitting plasma has a high density, over-abundant light metals, and a high velocity outflow—this combination of properties strongly suggests an association of GRB011211 with a very recent supernova, caused by the collapse of a massive progenitor star²⁶. In this model, the supernova ejects a substantial quantity of enriched material at high velocity ($\sim 0.1c$) into the surrounding medium, which is subsequently illuminated by the γ -ray burst. The high metal abundances of the illuminated material would arise naturally from the pre-supernova nucleosynthesis and the short interval between the supernova explosion and the burst itself. Indeed, the supernova model is theoretically preferred over neutron-star mergers as the explanation of long-duration γ -ray bursts⁷, as seen here. The supernova link is further strengthened by the observed correlation of long-duration bursts with star-forming regions^{27,28}, and, in one case, a direct association between the bright supernova 1998bw and the very-low-redshift ($z = 0.0086$) γ -ray burst, GRB980425²⁹.

Received 21 January; accepted 6 March 2002.

1. Metzger, M. R. *et al.* Spectral constraints on the redshift of the optical counterpart to the γ -ray burst of 8 May 1997. *Nature* **387**, 878–880 (1997).
2. Meszaros, P. Theories of gamma-ray bursts. Preprint astro-ph/0111170 at (<http://xxx.lanl.gov>) (2001).
3. Paczynski, B. Gamma-ray bursters at cosmological distances. *Astrophys. J.* **308**, L43–L46 (1986).
4. Eichler, D., Livio, M., Piran, T. & Schramm, D. N. Nucleosynthesis, neutrino bursts and γ -rays from coalescing neutron stars. *Nature* **340**, 126–128 (1989).
5. Mochkovitch, R., Hernanz, M., Isern, J. & Martin, X. Gamma-ray bursts as collimated jets from neutron star/black hole mergers. *Nature* **361**, 236–238 (1993).
6. Woosley, S. E. Gamma-ray bursts from stellar mass accretion disks around black holes. *Astrophys. J.* **405**, 273–277 (1993).
7. Paczynski, B. Are gamma ray bursts in star-forming regions? *Astrophys. J.* **494**, L45–L48 (1998).
8. Fryer, C. L., Woosley, S. E. & Hartmann, D. H. Formation rates of black hole accretion disk gamma-ray bursts. *Astrophys. J.* **526**, 152–177 (1999).
9. Frontera, F. *et al.* GRB011211: BeppoSAX/GRBM data. *GCN GRB Obs. Rep. No. 1215* (2001).
10. Fruchter, A. *et al.* GRB 011211: Spectroscopy of the optical counterpart. *GCN GRB Obs. Rep. No. 1200* (2001).
11. Holland, S. *et al.* The optical afterglow of the gamma-ray burst GRB 011211. *Astron. J.* (submitted); preprint astro-ph/0202309 at (<http://xxx.lanl.gov>) (2002).
12. Burud, I. *et al.* GRB 011211: Detection of the probable host galaxy. *GCN GRB Obs. Rep. No. 1213* (2001).
13. Jansen, F. *et al.* XMM-Newton observatory. I. The spacecraft and operations. *Astron. Astrophys.* **365**, L1–L6 (2001).
14. Santos-Lleo, M., Loiseau, N., Rodriguez, P., Altieri, B. & Scharrel, N. GRB 011211: XMM-Newton observation of GRB011211. *GCN GRB Obs. Rep. No. 1192* (2001).
15. Bevington, P. R. & Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences* (McGraw-Hill, New York, 1992).
16. Mewe, R., Gronenschild, E. H. B. M. & van den Oord, G. H. J. Calculated X-radiation from optically thin plasmas. *Astron. Astrophys. Suppl.* **62**, 197–254 (1985).
17. Ballantyne, D. R. & Ramirez-Ruiz, E. Iron K α emission from X-ray reflection: Predictions for gamma-ray burst models. *Astrophys. J.* **559**, L83–L87 (2001).
18. Kallman, T. R., Meszaros, P., Rees, M. J. Iron K lines from gamma ray bursts. *Astrophys. J.* (submitted); preprint astro-ph/0110654 at (<http://xxx.lanl.gov>) (2002).
19. Rees, M. J. & Meszaros, P. Fe K α emission from a decaying magnetar model of gamma-ray bursts. *Astrophys. J.* **545**, L73–L75 (2000).
20. Frail, D. A. *et al.* Beaming in gamma-ray bursts: Evidence for a standard energy reservoir. *Astrophys. J.* **562**, L55–L58 (2001).
21. Piro, L. *et al.* Observation of X-ray lines from a gamma-ray burst (GRB991216): Evidence of moving ejecta from the progenitor. *Science* **290**, 955–958 (2000).
22. Antonelli, L. A. *et al.* Discovery of a redshifted iron K line in the X-ray afterglow of GRB 000214. *Astrophys. J.* **545**, L39–L42 (2000).
23. Piro, L. *et al.* Iron line signatures in X-ray afterglows of GRB by BeppoSAX. *Astron. Astrophys. Suppl.* **138**, 431–432 (1999).
24. McLaughlin, G. C., Wijers, R. A. M. J., Brown, G. E. & Bethe, H. A. Broad and shifted iron-group emission lines in gamma-ray bursts as tests of the hypernova scenario. *Astrophys. J.* (submitted); preprint astro-ph/0110614 at (<http://xxx.lanl.gov>) (2001).
25. Weth, C., Meszaros, P., Kallman, T. & Rees, M. J. Early X-ray/ultraviolet line signatures of gamma-ray burst progenitors and hypernovae. *Astrophys. J.* **534**, 581–586 (2000).
26. Wheeler, J. C., Yi, I., Hoeflich, P. & Wang, L. Asymmetric supernovae, pulsars, magnetars, and gamma-ray bursts. *Astrophys. J.* **537**, 810–823 (2000).
27. Bulik, T., Belczynski, K. & Zbijewski, W. Distribution of compact object mergers around galaxies. *Mon. Not. R. Astron. Soc.* **309**, 629–635 (1999).
28. Hanlon, L. *et al.* ISO detection of a 60 μm source near GRB 970508. *Astron. Astrophys.* **359**, 941–947 (2000).
29. Bloom, J. S. *et al.* The unusual afterglow of the γ -ray burst of 26 March 1998 as evidence for a supernova connection. *Nature* **401**, 453–456 (1999).
30. Vietri, M. & Stella, L. Supranova events from spun-up neutron stars: an explosion in search of an observation. *Astrophys. J.* **527**, L43–L46 (1999).

Acknowledgements

This work is based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). We thank M. Rees, M. Davies, B. McBreen, R. Willingale and M. Barstow for critical reading of the manuscript, and for discussions.

Competing interests statement

The authors declare that they have no competing financial interests..

Correspondence and requests for materials should be addressed to J.N.R. (e-mail: jnr@star.le.ac.uk).

Bunching of fractionally charged quasiparticles tunnelling through high-potential barriers

E. Comfórti, Y. C. Chung, M. Heiblum, V. Umansky & D. Mahalu

Braun Center for Submicron Research, Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel

Shot noise measurements have been used to measure the charge of quasiparticles in the fractional quantum Hall (FQH) regime^{1–3}. To induce shot noise in an otherwise noiseless current of quasiparticles, a barrier is placed in its path to cause weak backscattering. The measured shot noise is proportional to the charge of the quasiparticles; for example, at filling factor $\nu = 1/3$, noise corresponding to $q = e/3$ appears. For increasingly opaque barriers, the measured charge increases monotonically, approaching $q = e$ asymptotically^{4,5}. It was therefore believed that only electrons, or alternatively, three bunched quasiparticles, can tunnel through high-potential barriers encountered by a noiseless current of quasiparticles. Here we investigate the interaction of $e/3$ quasiparticles with a strong barrier in FQH samples and find that bunching of quasiparticles in the strong backscattering limit depends on the average dilution of the quasiparticle current. For a very dilute current, bunching ceases altogether and the transferred charge approaches $q = e/3$. This surprising result demonstrates that quasiparticles can tunnel individually through high-potential barriers originally thought to be opaque for them.

In 1918, Schottky determined the charge of the electron by measuring the average square of the current fluctuations, $S = \langle i^2 \rangle$, resulting from the stochastic emission of electrons in a vacuum tube—naming it ‘shot noise’⁶. His expression $S = 2qI$ is a result of independent random events obeying Poisson distribution. Here, S is the spectral density of the fluctuations (in units of $A^2 \text{ Hz}^{-1}$), q is the charge of the particle and I is the average current. Similar experiments were implemented^{1–3} in the FQH regime⁷, verifying the existence of fractionally charged quasiparticles⁸. A partially transmitting constriction, a quantum point contact (QPC), served in these experiments as an adjustable potential barrier in the path of the current, thus partitioning the transmitted charges. This random process is described by a binomial distribution, resulting in $S = 2qI(1 - t)$ at zero temperature, with $t \in [0, 1]$ being the transmission coefficient of the QPC^{4,9,10}. In the weak backscattering regime, the quasiparticles were found to traverse the QPC independently of each other and the measured charge was $q = e/3$ at filling factor $\nu = 1/3$ and $q = e/5$ for $\nu = 2/5$ (refs 1,2,3). As backscattering gets stronger, the tunnelling of individual quasiparticles becomes correlated, and in the limit of a pinched QPC and

$\nu = 1/3$ three quasiparticles were found to group together to tunnel through the barrier. Obviously, Schottky’s formula is inapplicable for correlated (or bunched) quasiparticles, but it can still be used to characterize the system with the effect of bunching being incorporated into an effective charge $q(t)$. Hence, the noise for a pinched QPC becomes electronic-like⁴, that is, with an effective charge of $q = e$. Moreover, a nearly universal behaviour was found for the evolution of the effective charge $q(t)$ (ref. 5), starting at $q(\text{openQPC}) = e/3$ and monotonically increases toward $q(\text{pinchedQPC}) = e$.

Here we explore the bunching properties of a pinched QPC when a sparse beam of $e/3$ quasiparticles impinges upon it. In other words, when there are not enough quasiparticles in close proximity to bunch into an ‘electron’, we may ask the following questions: (1) will

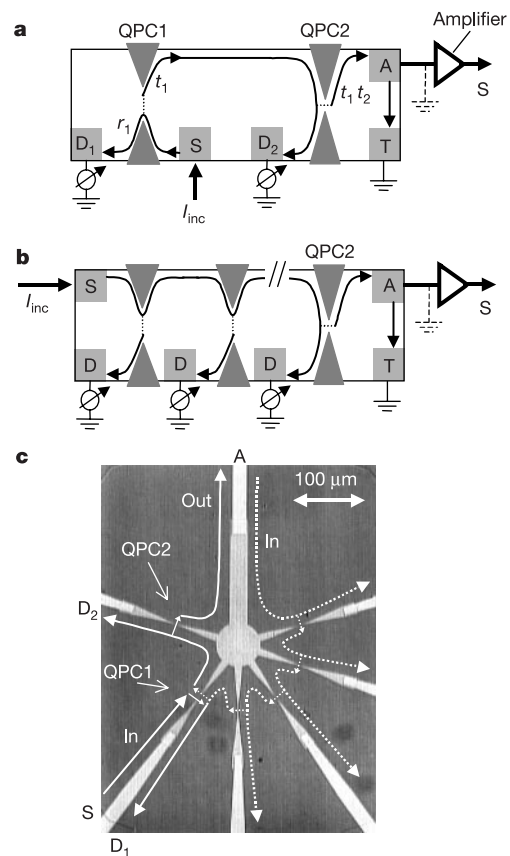


Figure 1 Schematic and actual representations of the quasiparticle injector followed by a quasiparticle filter, both made of quantum point contacts, QPC1 and QPC2, respectively. **a**, An injector made of a relatively open QPC1 partitions an incident noiseless (d.c.) current, injected from source terminal S. The scattered part (t_1), composed of a dilute beam of quasiparticles, impinges on a pinched QPC2. The resulting noise is measured by a cooled, low-noise, amplifier at terminal A (see ref. 1). The intermediate drain D_2 prohibits multiple reflections, and the grounded terminal T is used to fix the output impedance of the sample and make it independent of QPC settings. **b**, An alternative scheme, suitable for producing a moderately dilute current, invokes a cascade of weakly backscattering QPCs transmitting a dilute quasiparticle beam (see ref. 11). **c**, A photograph of the actual device in the vicinity of the QPCs, formed by metallic gates (light grey regions) deposited on the surface of the GaAs–AlGaAs heterostructure, embedding a two-dimensional electron gas some $0.1 \mu\text{m}$ below the surface. Electron mobility is $2 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ and areal carrier density is $1.1 \times 10^{11} \text{ cm}^{-2}$, both measured at 4.2 K in the dark. The solid arrows correspond to the direction of current in configuration **a**, and the other QPCs on the right (with current flow denoted by dotted arrows) are used when configuration **b** is employed. Ohmic contacts (serving as S, D, T, A) are outside the frame of the picture.

the barrier become opaque? (2) Will quasiparticles be 'borrowed' from lower-energy states to fill in for the missing ones in the beam? Or, (3) perhaps the bunching condition will be relaxed and individual quasiparticles will transverse the barrier? Theory does not yet provide answers to these questions.

Samples were fabricated in a high-mobility two-dimensional electron gas embedded in a GaAs–AlGaAs heterostructure. Measurements were done in a strong magnetic field of $B \approx 13$ T and a fractional filling factor of $\nu = 1/3$ in the FQH regime. Vanishing of the longitudinal resistivity ρ_{xx} ensures that the (net) current is flowing chirally along the edges of the sample in edge states. This allows measurements in multi-terminal geometries, shown in Fig. 1. Two techniques are used to partition the quasiparticle beam, hence making it sparse or dilute, before it impinges on the pinched QPC2. A straightforward scheme, shown in Fig. 1a, uses a noiseless current I_{inc} that impinges on a relatively open QPC1 and partially scatters toward QPC2 (although in this case the small reflection coefficient of QPC1 is responsible for the dilution of the current, for uniformity we stick to the notation of transmission coefficient $t_1 \rightarrow 0$). Most of the current continues toward drain D₁; the scattered part toward QPC2 is a very dilute beam of quasiparticles with charge $q_1 = e/3$ (refs 1, 2, 5) and dilution determined by t_1 . Much of that dilute current is reflected back by QPC2 toward drain D₂ and a small part t_2 is transmitted. A fraction $t_1 t_2$ of the incident current reaches the amplifier. This method cannot be applied, however, to achieve a moderately sparse beam of quasiparticles since a partly pinched QPC1 would lead to bunching of quasiparticles and to an effective charge of $q_1 > e/3$ (ref. 5). Hence, we also use a geometry shown in Fig. 1b. Here, the incident current is being partitioned by transmission through a cascade of weakly backscattering QPCs (for each $t_1 \rightarrow 1$), feeding QPC2 with a current

of quasiparticles with arbitrary dilution $t_1 = \Pi_i t_i$. The fact that the partitioned charge by this method is $e/3$ is not obvious and a detailed study¹¹ was needed to prove that a beam of quasiparticles is indeed produced. We also tested, through detailed noise measurements¹¹, whether dilute quasiparticles suffer intra-edge scattering and subsequent equilibration during transport along the device edges. Equilibration establishes a new chemical potential and increases the occupation of each state below the chemical potential, hence modifying the dilution of the beam. As shown in ref. 11, such equilibration does not take place in our devices.

Using one of the two methods depicted in Fig. 1, we create a noisy beam of quasiparticles with charge $e/3$, which is being further partitioned by QPC2. The noise at amplifier A is measured with a spectrum analyser after amplification by a cooled amplifier. The amplifier, being placed near the sample, has a very low current noise at its input, $\langle i_{\text{amp}}^2 \rangle = 1.5 \times 10^{-28} \text{ A}^2 \text{ Hz}^{-1}$, when operating with a bandwidth of 30 kHz and centre frequency $f_0 \approx 1.5$ MHz. The value of f_0 , chosen well above the cut-off of the ubiquitous $1/f$ noise, is determined by a resonance of LC circuit, with C the capacitance of the coaxial cable connecting the sample and the amplifier and L the inductance of an added superconducting coil¹. Reflected currents flow into the grounded terminals D and T, leading to a constant input (at S) and output (at A) conductance $G = e^2/3h$ that is independent of the transmission of the QPCs. This makes both the sample's equilibrium noise ($4k_B T G = 5 \times 10^{-29} \text{ A}^2 \text{ Hz}^{-1}$) and the sample-dependent amplifier's noise ($\langle i_{\text{amp}}^2 \rangle / G^2$) independent of the transmission of the QPCs, allowing subtraction of both from the measured noise (for comparison, the magnitude of the shot noise at A is typically in the $10^{-30} \text{ A}^2 \text{ Hz}^{-1}$ range). k_B is Boltzmann's constant.

The configuration in Fig. 1a can be analysed by means of superposition¹². Consider first the 'injector' QPC1, characterized by transmission $t_1 \rightarrow 0$ toward QPC2 and partitioned charge $e/3$. Because QPC1 is a stochastic element, it generates (at zero tem-

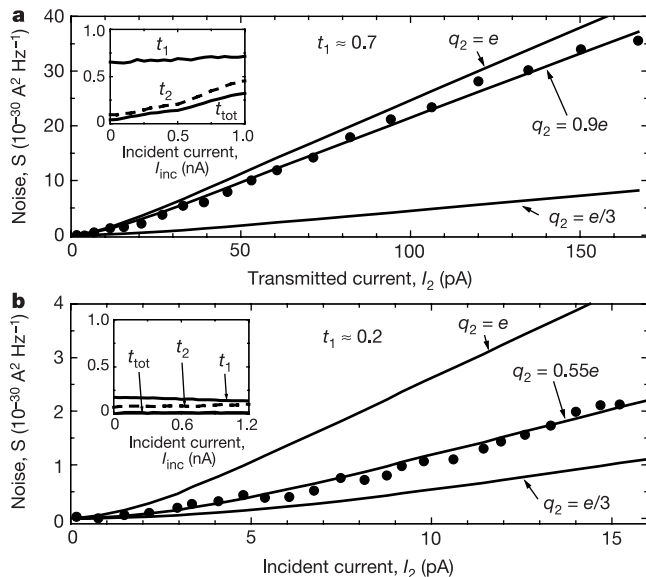


Figure 2 Noise and transmission measurements of the pinched QPC2 (with transmission $t_2 \approx 0.1$ at zero bias) for two different values of dilution of the impinging current: $t_1 \approx 0.7$ (a) and 0.2 (b). In the main graphs the measured noise is plotted against the transmitted current, together with the theoretical prediction of the independent particles model at a finite temperature. The intermediate curve in each graph represents the best fit to an arbitrary charge q_2 . Various transmission coefficients, measured simultaneously with the noise, are shown in the insets against the incident (noiseless) current. Each inset shows the dilution level t_1 generated by the QPC1 injector, the transmission t_2 of the pinched QPC2 in response to the dilute impinging current, and the total transmission t_{tot} . We note that the sensitivity of t_2 to the current depends on the dilution level of the impinging current.

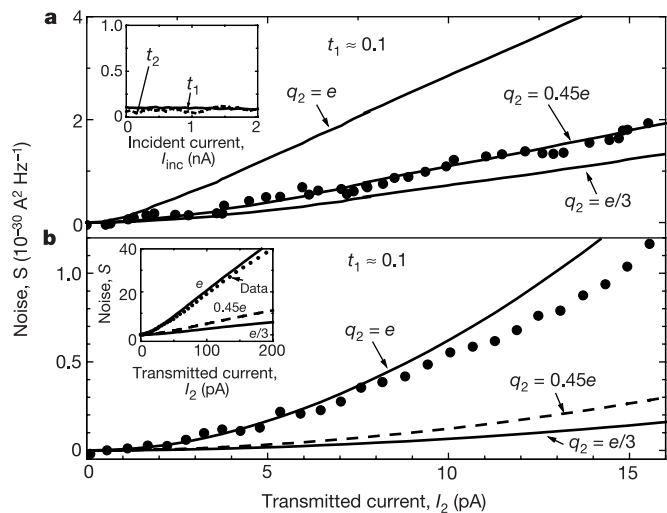


Figure 3 Comparison of the charge characterizing the pinched QPC2 for two extreme cases of the impinging current: not diluted (noiseless) and highly dilute, keeping the same transmitted current. **a**, The noise produced by the pinched QPC2 when fed by a highly dilute impinging current, $t_1 \approx 0.1$, corresponds to a quasiparticle charge $q_2 = 0.45e$. The inset shows the current-dependent transmission t_1 (level of dilution) and the transmission t_2 (which is fairly current-independent for such a dilute beam). **b**, The noise produced by the pinched QPC2 when fed by a noiseless current corresponds to an almost electronic charge. The inset verifies the charge $q_2 = e$ by measuring the noise over a considerably wider range of transmitted current. Clearly, the charge characterizing QPC2 depends not only on the potential barrier height but also on the average occupation of the states (dilution) of the impinging current.

perature) noise $2(e/3)I_{\text{inc}}t_1(1 - t_1)$, with $I_{\text{inc}}t_1$ the transmitted current, impinging on QPC2. This noise is attenuated with a factor t_2^2 by QPC2, resulting in a contribution of QPC1 to the total noise:

$$S_1 = 2(e/3)I_{\text{inc}}t_1(1 - t_1)t_2^2 \quad (1)$$

Consider now QPC2, characterized by transmission t_2 and charge q_2 when impinged by a noiseless current I_{imp} of $e/3$ quasiparticles. It produces noise $2q_2I_{\text{imp}}t_2(1 - \tilde{t}_2)$, where $I_{\text{imp}}t_2 = I_2$ is the transmitted current and $\tilde{t}_2 = t_2[(e/3)/q_2]$ denotes the effective transmission for charge q_2 quasiparticles. This transmission $\tilde{t}_2$ is determined self-consistently with the charge q_2 in order to maintain the measured conductance of QPC2 (ref. 5). We stress that even though the current $I_{\text{imp}} = I_{\text{inc}}t_1$ is noisy we still use the above expression to calculate the noise generated by QPC2, because the noise in I_{imp} was already taken into account in equation (1). The added contribution of QPC2 is therefore:

$$S_2 = 2q_2I_{\text{inc}}t_1t_2(1 - \tilde{t}_2) \quad (2)$$

The total noise in A is then $S_1 + S_2$. The correctness of this analysis can be validated in the limit of a constant charge (say, $e/3$): $S_1 + S_2 = 2(e/3)I_{\text{inc}}t_1t_2(1 - t_1t_2)$, with $t_{\text{tot}} = t_1t_2$ being the total transmission from S to A—the standard expression for a binomially distributed process. In the experiment we use the expression for $S_1 + S_2$ in order to determine the charge q_2 partitioned by QPC2. In the limit where both t_1 and t_2 are small, S_1 and S_2 are of the order of $O(t_1t_2^2)$ and $O(t_1t_2)$, respectively, with the S_1 much smaller than S_2 . Hence, the measured noise is dominated by the contribution of the pinched QPC2:

$$S \equiv 2q_2I_{\text{inc}}t_1t_2 = 2q_2I_2 \quad (3)$$

We verify first that the noise produced by a pinched QPC, when fed with a quiet current, corresponds as expected to $q \approx e$. We find results similar to these in ref. 5 with an example given in Fig. 3b. For $t_1 = 1$, hence feeding QPC2 with a noiseless current, and $t_2 \approx 0.1$, we measure indeed charge e . We then partition the incident current by setting $t_1 < 1$, and hence impinging a noisy current of quasiparticles on QPC2 with $t_2 \approx 0.1$. Two examples are shown in Fig. 2, one for average state occupation $t_1 \approx 0.7$ (Fig. 2a) and one for $t_1 \approx 0.2$ (Fig. 2b). When calculating the expected noise⁵ we take into

account the finite temperature of the electrons ($T \approx 65$ mK) and the energy (or current) dependence of the total transmission $t_{\text{tot}} = t_1t_2$ (current-dependent transmissions are shown in the insets of Fig. 2). The average current is being varied over a large enough range, with the voltage V satisfying $q_2V \gg k_B T$, to allow the noise to reach the linear regime. Nice agreement is found between the data and the independent particle model for $q_2 = 0.9e$ and $t_1 \approx 0.7$ (lightly diluted current) and for $q_2 = 0.55e$ and $t_1 \approx 0.2$ (highly diluted current).

A more striking example of the effect of beam dilution is demonstrated in Fig. 3, where the range of current I_2 is kept constant for different values of dilution. Obviously, a higher source voltage is required to obtain the same current I_2 when the current is more dilute. In comparison with the measured electron charge for a noiseless impinging current (Fig. 3b), a highly dilute current ($t_1 \approx 0.1$) impinging on the pinched QPC2 is found to produce a small charge $q_2 = 0.45e$ (Fig. 3a)—slightly above the charge of the quasiparticles.

Figure 4 summarizes the dependence of q_2 on the dilution t_1 of the impinging current on QPC2. Two examples, $t_2 = 0.1$ and $t_2 = 0.25$, are chosen, with corresponding charge e and $0.75e$, respectively, for a noiseless impinging current. The more dilute the impinging current is ($t_1 \rightarrow 0$), the smaller is the effective charge q_2 —approaching $\sim e/3$ asymptotically. We conclude that bunching of quasiparticles is not an essential mechanism for the transfer of quasiparticles through high-potential barriers. Bunching takes place only when the incoming states are highly occupied.

We also wondered how the transmission of QPC2 is affected by the dilution of the impinging quasiparticles. Present theory assumes only a noiseless current approaching a constriction within the framework of the Luttinger model. Here we also find counter-intuitive results. In the linear regime, where the source voltage is small enough to keep the transmission almost energy-independent (Fig. 5a), the transmission t_2 is independent of dilution (although the source voltage for the dilute current is some ten times larger). This can be compared with the case where the same source-voltage range is kept (Fig. 5b). Here, the transmission t_2 of the noiseless current strongly depends on voltage (approaching unity at $V > 50 \mu\text{V}$). Equilibration of quasiparticles had been ruled out¹¹,

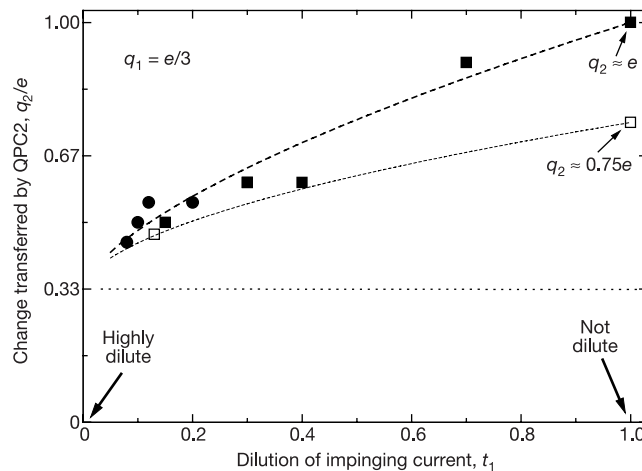


Figure 4 Evolution of the effective charge q_2 that characterizes the pinched QPC2 in response to different values of dilution t_1 of the impinging current (extracted from curves similar to that in Fig. 3a). Results of three different measurements are shown—two complementary sets of data for $t_2 = 0.1$ (filled squares and circles), with dilution produced by backscattering of a single QPC1 (circles, Fig. 1a) and by transmission through five, relatively open, QPCs (squares, Fig. 1b), and one set with $t_2 = 0.25$. In the case of $t_2 = 0.25$ only the extreme points are shown in order to simplify the graph (open

squares). The dashed lines are only a guide to the eye. Evidently, as the current impinging on the pinched QPC2 becomes more dilute, the charge drops from its original value toward $\sim e/3$ in the limit of very high dilution. We conclude that individual, very sparse, quasiparticles tunnel through a pinched QPC that was originally thought to be opaque for them. Filled circles and squares, $t_2 = 0.1$ and $q_2(\text{notdilute}) = e$. Open squares, $t_2 = 0.25$ and $q_2(\text{notdilute}) = 0.75e$.

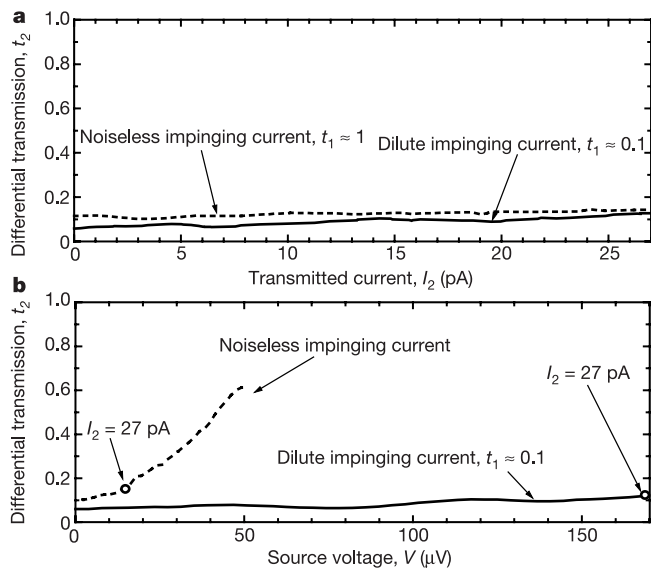


Figure 5 Dependence of the transmission t_2 of the pinched QPC2 on the dilution t_1 of the impinging current. **a**, Measurements in the linear regime. The transmission t_2 of a highly dilute impinging current ($t_1 \approx 0.1$, solid curve) and that of a noiseless current ($t_1 = 1$, dashed curve), with the same transmitted current range kept in both cases. The applied source voltage, on the other hand, reaches a maximal value of ~ 170 μ V for the noisy impinging current but only ~ 16 μ V in the noiseless case. Nevertheless, the transmissions in both cases are similar. **b**, Measurements in the nonlinear regime. Similar measurements to **a** but the same applied source voltage range is kept in both cases (the noiseless current is some ten times larger for the same voltage). The transmission is found to be strongly current-dependent when the impinging current is noiseless.

so we conclude that the nonlinearity of the pinched QPC depends strongly on the quasiparticle current and less on the quasiparticle energy. This rules out that the potential profile of the barrier in the QPC is responsible for the nonlinearity of the current. Moreover, the insensitivity of the transmission t_2 in the linear regime to the dilution of the impinging quasiparticle beam suggests equal probabilities of tunnelling for a single quasiparticle and for bunched quasiparticles. In other words, noise and transmission measurements show that quasiparticles can transfer, with the same ease, either one by one or bunched in groups. Their bunching depends on the transparency of the barrier and on the preparation of the quasiparticle beam. It is now for theory to explain such a bizarre effect.

Received 6 November 2001; accepted 20 February 2002.

- de Picciotto, R. *et al.* Direct observation of a fractional charge. *Nature* **389**, 162–164 (1997).
- Saminadayar, L., Glatli, D. C., Jin, Y., Etienne, B. Observation of the $e/3$ fractionally charged Laughlin quasiparticle. *Phys. Rev. Lett.* **79**, 2526–2529 (1997).
- Reznikov, M., de-Picciotto, R., Griffiths, T. G., Heiblum, M. & Umansky, V. Observation of quasiparticles with one-fifth of an electron's charge. *Nature* **399**, 238–241 (1999).
- Kane, C. L. & Fisher, M. P. A. Nonequilibrium noise and fractional charge in the quantum Hall effect. *Phys. Rev. Lett.* **72**, 724–727 (1994).
- Griffiths, T. G., Comforti, E., Heiblum, M., Stern, A. & Umansky, V. Evolution of quasiparticle charge in the fractional quantum hall regime. *Phys. Rev. Lett.* **85**, 3918–3921 (2000).
- Schottky, W. Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern. *Ann. Phys. (Leipzig)* **57**, 541–567 (1918).
- Prange, R. E. & Girvin, S. M. (eds) *The Quantum Hall Effect* (Springer, New York, 1987).
- Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractional charge excitations. *Phys. Rev. Lett.* **50**, 1395–1398 (1982).
- Khlus, V. K. Current and voltage fluctuations in microjunctions of normal and superconducting metals. *Sov. Phys. JETP* **66**, 1243–1249 (1987).
- Lesovik, G. B. Excess quantum noise in 2D ballistic point contacts. *JETP Lett.* **49**, 592–594 (1989).
- Comforti, E., Chung, Y. C., Heiblum, M. & Umansky, V. Multiple scattering of fractionally-charged quasiparticles. Preprint cond-mat/0112367 at (<http://xxx.lanl.gov>) (2001).
- Oliver, W. D., Kim, J., Liu, R. C. & Yamamoto, Y. Hanbury Brown and Twiss-type experiment with electrons. *Science* **284**, 299–301 (1999).

Acknowledgements

The work was partly supported by the Israeli Academy of Science and by the German-Israeli Foundation (GIF). We thank A. Yacoby, A. Stern and Y. Levinson for discussions.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.H. (e-mail: heiblum@wisemail.weizmann.ac.il).

Atomic-scale images of charge ordering in a mixed-valence manganite

Ch. Renner*, G. Aeppli*, B.-G. Kim†, Yeong-Ah Soh* & S.-W. Cheong†

* NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540, USA

† Rutgers University, Department of Physics and Astronomy, Piscataway, New Jersey 08854, USA

Transition-metal perovskite oxides exhibit a wide range of extraordinary but imperfectly understood phenomena. The best known examples are high-temperature superconductivity in copper oxides¹, and colossal magnetoresistance in manganese oxides ('manganites')^{2,3}. All of these materials undergo a range of order–disorder transitions associated with changes in charge, spin, orbital and lattice degrees of freedom. Measurements of such order are usually made by diffraction techniques, which detect the ionic cores and the spins of the conduction electrons. Unfortunately, because such techniques are only weakly sensitive to valence electrons and yield superpositions of signals from distinct submicrometre-scale phases, they cannot directly image phase coexistence and charge ordering, two key features of the manganites. Here we present scanning tunnelling microscope measurements of the manganite $\text{Bi}_{1-x}\text{Ca}_x\text{MnO}_3$. We show that charge ordering and phase separation can be resolved in real space with atomic-scale resolution. By taking together images and current–voltage spectroscopy data we find that charge order correlates with both structural order and the local conductive state (either metallic or insulating). These experiments provide an atomic-scale basis for descriptions⁴ of manganites as mixtures of electronically and structurally distinct phases.

The material chosen for our experiments is $\text{Bi}_{1-x}\text{Ca}_x\text{MnO}_3$ (BCMO). For trivalent Bi and divalent Ca, the Mn ions are in a mixed valence state, Mn^{3+x} . At high temperature, Mn^{3+} and Mn^{4+} randomly occupy the manganese sites. Upon reducing the temperature, these cations are believed to order, yielding an increased lattice periodicity visible to X-ray and neutron diffraction⁵. For our samples of nominal $x = 0.76$, grown from a BiO flux, this occurs at $T_{\text{CO}} = 250$ K, as established using SQUID (superconducting quantum interference device) magnetometry. We performed the scanning tunnelling microscope (STM) experiments in ultrahigh vacuum at a base pressure of 5×10^{-10} torr. Previously published STM investigations of manganites primarily focused on spectroscopy of the density of states averaged over many atoms^{6,7}, and demonstrated phase separation into metallic and insulating regions on submicrometre⁸, but not atomic, length scales. In contrast, STM with atomic resolution has already been achieved for copper oxides, and revealed inhomogeneities in the superconducting order on atomic length scales^{9,10}.

Single crystals of BCMO do not cleave naturally, and preparing

flat and atomically clean surfaces suitable for STM is a challenge. We found that cleaning an as-grown sample in ethanol shortly before loading it into the vacuum chamber allowed us to obtain reproducible topographic images with atomic resolution. The STM tips were made of etched tungsten wires. Typical tunnelling parameters during imaging were 0.2 nA and 0.7 V for the set current and sample bias, respectively. The piezoelectric scan actuators were calibrated to an accuracy of 4% by imaging layered graphite crystals at different temperatures.

As might be expected from the rather crude surface preparation, STM failed to yield atomic resolution over large portions of the surface. We mainly saw a broad range of different nanometre-scale textures. However, we repeatedly found clean terraces with areas of a few hundred nm² where atomic resolution was routinely achieved. The room-temperature image in Fig. 1a clearly shows the square lattice expected for the (010) face of BCMO (simple cubic notation). The corresponding lattice constant $a_0 = 3.8 \pm 0.6 \text{ \AA}$ is in good agreement with the value of $3.77 \text{ \AA}$ determined by X-ray diffraction^{11,12}. At room temperature, we sometimes also observed a ' $\sqrt{2}a_0 \times \sqrt{2}a_0$ ' lattice, with the unit cell doubled along [101], coexisting with the ordinary square lattice (Fig. 2a). In Fig. 2b we show a section (shown orange in Fig. 2a) taken along one crystallographic axis. This section suggests a homogeneous charge distribution among the atomic sites in the cubic (disordered) region, whereas charges appear redistributed among alternating atomic sites in the ordered region, generating the doubled unit cell. The most natural interpretation is that there is local charge ordering—namely, the alternation of Mn³⁺ and Mn⁴⁺ ions—as posited in the simplest descriptions of mixed-valence manganites. Conductance spectra acquired in all areas where the doubled unit cell is stable systematically show more insulating characteristics than those measured in cubic regions (Fig. 2c). This validates the association of charge ordering with a metal–insulator transition—the carriers are localized and the material is insulating when charges are distributed in a set spatial pattern rather than fluctuating with time at each site.

The metallic state is not an especially good metal, with a semimetallic spectrum superposed on a small ohmic term (Fig. 2c

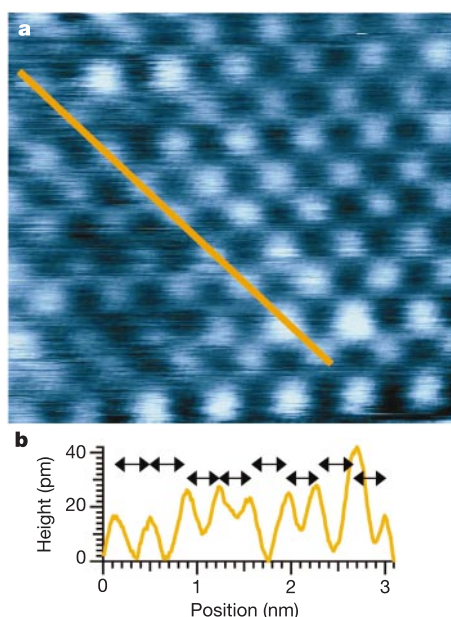


Figure 1 STM mapping of the paramagnetic phase of Bi_{0.24}Ca_{0.76}MnO₃ at 299 K. **a**, $3.4 \times 3.1 \text{ nm}^2$ image with a well-resolved square lattice ($a_0 = 3.8 \pm 0.6 \text{ \AA}$). **b**, The intensity profile along the line shown orange in **a** reveals a random distribution of short and long interatomic distances (indicated by arrows) along a main crystallographic axis.

inset). The energy at which the deviations from ohmic behaviour become apparent is around 0.1 eV, which is also where a strong non-Drude contribution begins to dominate the optical conductivity for similar samples at room temperature¹³. The gap for the insulating spectrum collected from the regions at room temperature where the doubled unit cell occurs is $\sim 0.7 \text{ eV}$. This is very similar to the gap that can be deduced from the optical conductivity measured at 150 K, when the bulk of the crystal is charge-ordered¹³. At a qualitative level, the spectra of Fig. 2c look remarkably similar to those measured by Fäth *et al.*⁸ on the isostructural compound (La,Ca)MnO₃. Although their lower hole-doped samples lack a charge-ordering phase transition, their finding of spectroscopic inhomogeneities near the metal–insulator transition is consistent with our results. Figure 2 shows the first data (to our knowledge) to

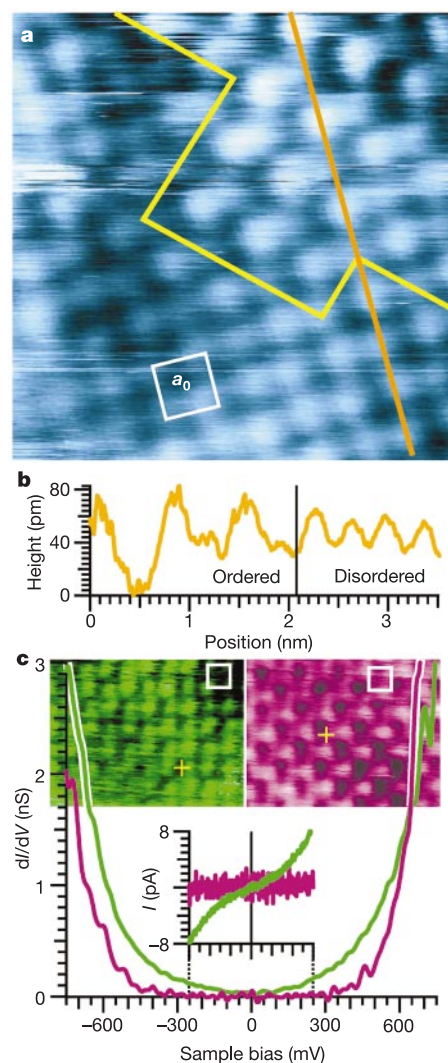


Figure 2 Topographic and spectroscopic atomic-scale signatures of phase separation into metallic and insulating regions in the paramagnetic phase of Bi_{0.24}Ca_{0.76}MnO₃ at 299 K. **a**, $3.5 \times 3.5 \text{ nm}^2$ STM image of a grain boundary (yellow line) between an insulating $\sqrt{2}a_0 \times \sqrt{2}a_0$ charge-ordered region (upper right) and a more metallic homogeneous cubic region (lower left). **b**, Intensity profile extracted along the orange line in **a**. Note the larger amplitude modulation in the ordered region owing to charge ordering. **c**, Charge-ordered regions with the $\sqrt{2}a_0 \times \sqrt{2}a_0$ lattice (purple) yield insulating $dI/dV(V)$ characteristics, while the disordered cubic regions (green) are characterized by more metallic $dI/dV(V)$ characteristics (numerical derivatives normalized to the metallic junction resistance $R = V/I$ at 0.7 V). The low-bias part of the corresponding $I(V)$ data are shown in the inset. The spectra were taken at the yellow crosses on the $3.7 \times 2.9 \text{ nm}^2$ STM images (white squares, cubic unit cell).

connect the different tunnelling spectra to distinct phases at the atomic scale.

When the sample is cooled below T_{CO} , the $\sqrt{2}a_0 \times \sqrt{2}a_0$ lattice becomes the dominant structure (Fig. 3). The current-voltage spectra of these regions always show insulating characteristics notably similar to those measured in the ordered regions above T_{CO} , as illustrated in Fig. 4. Figure 3a displays a large region of uninterrupted $\sqrt{2}a_0 \times \sqrt{2}a_0$ order of precisely the type seen at the upper right of Fig. 2a. There is a clear modulation of the height as well as the positions of the atoms away from the vertices of a square lattice. A more thorough analysis of Fig. 3a reveals a finer structure, characterized by alternating long and short interatomic distances along the main crystallographic axes (Fig. 3b). The short interatomic distances form regular zigzag chains across the entire scanned area (yellow lines in Fig. 3a and d), thus breaking the crystal symmetry. The distortion amplitude, which is the difference between long and short interatomic distances, can be inferred directly from histograms of all interatomic distances measured along the cubic unit cell vectors (Fig. 3c).

This distortion, generally in excess of $0.5 \text{ \AA}$, is too large to ascribe to the Mn atoms alone. On the other hand, the bright spots in STM images such as Fig. 3a match the apical oxygens sitting over the manganese sites on the (010) surface according to X-ray experiments¹¹ (Fig. 3d). The X-ray data¹¹ show that the MnO_6 octahedra are tilted by an angle ν of about 10° , hence shifting the apical oxygen by a distance $d_{||} = 0.34 \text{ \AA}$ away from the vertices of the square lattice, where $d_{||} = a_{MO}\sin(\nu)$, and $a_{MO} = 1.94 \text{ \AA}$ is the Mn–O bond length. The tilting is not uniform: rather, the octahedra

form a zigzag pattern where neighbouring oxygens are either separated or brought closer by roughly $2d_{||}$. For a lattice constant of $3.77 \text{ \AA}$, these shifts should yield a lattice with alternating interatomic distances of $3.1 \text{ \AA}$ and $4.4 \text{ \AA}$, in excellent agreement with our STM images (Fig. 3c). We note that the room-temperature images bear a similar distribution of short and long interatomic distances. However, in contrast to the low-temperature images, the interatomic distances—short or long—are more randomly distributed (Fig. 1b). Static disorder in the oxygen tilts^{14,15} therefore appears to be annealed on cooling through T_{CO} , yielding a regular zigzag pattern. Although the precise temperature at which this happens at the surface has not been identified in the present experiments, these results illustrate on a local scale the idea that ordered lattice distortions, especially tilts of the MnO_6 octahedra, are important for stabilizing the charge-ordered state in manganites¹⁶.

Figure 3a contains not only information about the geometry of the crystal face exposed, but also about the ‘heights’ of its features. It is natural to ask whether the apparent ‘height’ modulation is derived simply from a modulation of atomic coordinates (possible surface reconstruction), or whether it indicates a modulation of the electronic wavefunctions centred at the atom cores. The above analysis of the in-plane displacements shows that the STM contrast consistently reflects the tilting of the MnO_6 octahedra. The apex oxygens illustrated as the larger blue spheres in Fig. 3d move a certain distance perpendicular to the (010) plane when the octahedra distort and tilt due to the Jahn–Teller effect induced by a central Mn^{3+} . Because all octahedra are tilted by approximately the same angle ($\nu \approx 10^\circ$), the apex oxygens are all displaced by roughly the

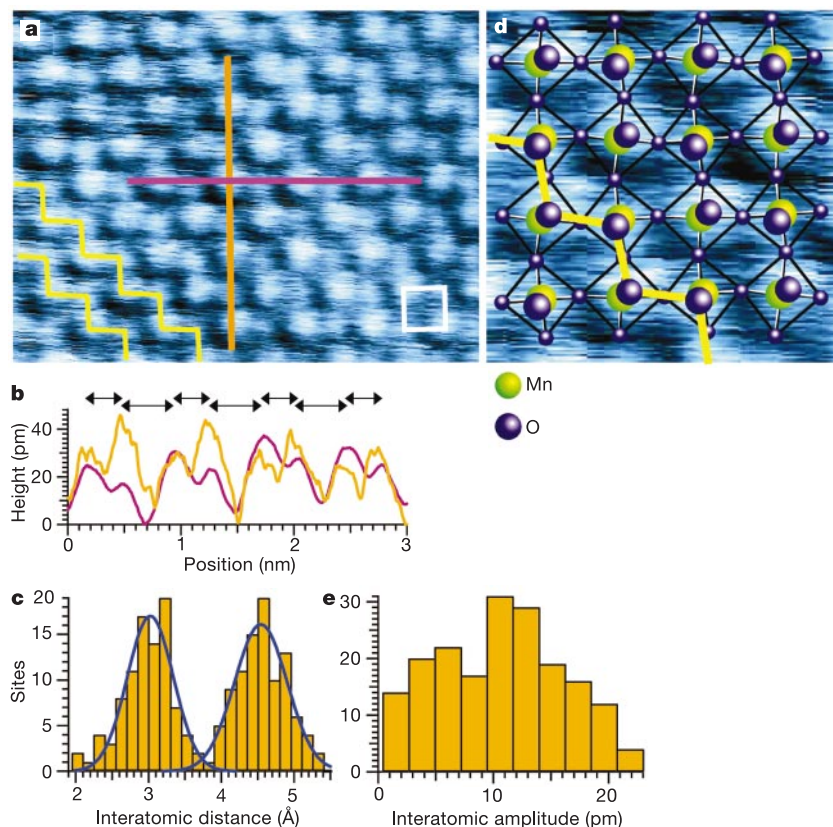


Figure 3 Atomic-scale STM mapping of the charge-ordered phase of $\text{Bi}_{0.24}\text{Ca}_{0.76}\text{MnO}_3$ at 146 K . **a**, $4.8 \times 3.6 \text{ nm}^2$ image of the $\sqrt{2}a_0 \times \sqrt{2}a_0$ lattice (white square, cubic unit cell). **b**, Intensity profiles extracted along the main crystallographic directions in **a** (orange and purple lines). They show the distortion of the atomic positions away from the cubic vertices

with alternating short and long interatomic distances (arrows). **c**, Bimodal distribution into short ($3.0 \pm 0.2 \text{ \AA}$) and long ($4.5 \pm 0.3 \text{ \AA}$) interatomic distances along $[100]$ and $[001]$. **d**, X-ray refined crystal structure of the (010) plane of the isostructural compound $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$ ($x = 1/2$) adapted from ref. 11 superposed on a magnified area of **a** ($1.7 \times 1.7 \text{ nm}^2$). For purposes of illustration, the apical oxygens are sketched larger than the in-plane oxygen. **e**, The amplitude difference between neighbouring atomic sites often

† Present address: University of the Faroe Islands, PB 2109, FO-165 Argir, Faroe Islands.

same distance $d_{\perp} = a_{\text{MO}}(1 - \cos(\nu))$. Thus, the tilting of the octahedra cannot account for the observed corrugation. On the other hand, the Jahn–Teller distortion of the order 0.05–0.1 Å (refs 11 and 17) happens exclusively for octahedra centred at Mn^{3+} . Taking this together with the fact that the octahedra are tilted, the Jahn–Teller distortion will then account for at most 0.1 Å of the height modulation at the apical oxygen sites. What we see instead—as demonstrated by the histogram in Fig. 3e—are apparent height differences between neighbouring atomic sites which are frequently larger than 0.1 Å. Thus, the measured corrugation needs an additional contribution, which is probably due to different wavefunctions responsible for the tunnelling between the two neighbouring octahedra and the STM tip. The most natural explanation is that the Mn ions at the centres of these octahedra have different charges and therefore different orbital occupancy, which affects the tunnelling through the apical oxygens that cap the planes exposed to our STM experiments.

The charge-ordered STM images discussed so far consistently show a Mn^{3+} to Mn^{4+} ratio of 1:1, in apparent contradiction with diffraction data^{5,12} (which show a periodicity of $\sim\sqrt{2}a_0$ rather than the $\sqrt{2}a_0$ that we see) for crystals prepared following the same procedures and with the same nominal doping. There are many possible explanations for this discrepancy. The first is to appeal to the surface sensitivity of STM, and assert that the surfaces behave differently from the bulk. Indeed, such an argument has been made for Sr_2RuO_4 , the only other transition-metal oxide without copper for which atomic-resolution STM data have been reported¹⁸. The agreement of our atomic coordinates with those deduced from bulk X-ray crystallography (Fig. 3d) speaks against a surface reconstruction on the scale of that seen for Sr_2RuO_4 . Furthermore, the agreement of the tunnelling gap shown in Fig. 4 with the optical gap¹³ (another bulk probe) also speaks against a surface effect. This leaves the possibility that even though the bulk charge-ordering transition for our crystals is as sharp as any previously reported, the effective doping may not be $x = 0.76$ at all surfaces. Besides the surface, which may affect the actual local doping, there is the

possibility of phase separation into (commensurate) charge-ordered phases upon cooling through T_{CO} (ref. 19). It is worth noting that what distinguishes our experiments from previous STM work on manganites is that we have atomic resolution for a small fraction, rather than none, of the crystal surface exposed. This leaves the possibility that the remaining fraction could exhibit structures consistent with bulk diffraction data for the nominal composition of the sample. The fact that the measured optical gap¹³ is marked by a smooth rise in conductivity rather than a hard onset is consistent with coexistence, even in the bulk, of distinct ordered phases with different periodicities and different gaps.

The atomic-scale images that we report here display many of the phenomena that have been posited for the manganites, most notably charge ordering, of which STM is the most direct probe because of its sensitivity to the outer valence electrons—diffraction and transmission electron microscopy, which have been the tools for the initial exploration of the charge-ordering phenomena, are all primarily sensitive to atomic core positions rather than the outer electrons. We have been able to associate the metallic and insulating current–voltage characteristics with distinct atomic-scale structures.

Received 17 December 2001; accepted 25 February 2002.

- Orenstein, J. & Millis, A. J. Advances in the physics of high-temperature superconductivity. *Science* **288**, 468–474 (2000).
- von Helmolt, R., Wecker, J., Holzapfel, B., Schultz, L. & Samwer, K. Giant negative magnetoresistance in perovskite-like $\text{La}_{2/3}\text{Ba}_{1/3}\text{MnO}_3$ ferromagnetic films. *Phys. Rev. Lett.* **71**, 2331–2333 (1993).
- Jin, S. *et al.* Thousand-fold change in resistivity in magnetoresistive La–Ca–Mn–O films. *Science* **264**, 413–415 (1994).
- Moreo, A., Yunoki, S. & Dagotto, E. Phase separation scenario for manganese oxides and related materials. *Science* **283**, 2034–2040 (1999).
- Bao, W., Axe, J. D., Chen, C. H. & Cheong, S.-W. Impact of charge ordering on magnetic correlations in perovskite $(\text{Bi,Ca})\text{MnO}_3$. *Phys. Rev. Lett.* **78**, 543–546 (1997).
- Wei, J. Y. T., Yeh, N.-C. & Vasquez, R. P. Tunneling evidence of half-metallic ferromagnetism in $\text{La}_{0.7}\text{Ca}_{0.3}\text{MnO}_3$. *Phys. Rev. Lett.* **79**, 5150–5153 (1997).
- Biswas, A., Elizabeth, S., Raychaudhuri, A. K. & Bhat, H. L. Density of states of hole-doped manganites: A scanning-tunneling-microscopy/spectroscopy study. *Phys. Rev. B* **59**, 5368–5376 (1999).
- Fäth, M. *et al.* Spatially inhomogeneous metal-insulator transition in doped manganites. *Science* **285**, 1540–1542 (1999).
- Howald, C., Fournier, P. & Kapitlnik, A. Inherent inhomogeneities in tunneling spectra of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8-x}$ crystals in the superconducting state. *Phys. Rev. B* **64**, R100504-1–R100504-4 (2001).
- Pan, S. H. *et al.* Microscopic electronic inhomogeneity in the high- T_c superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8-x}$. *Nature* **413**, 282–285 (2001).
- Radaelli, P. G., Cox, D. E., Marezio, M. & Cheong, S.-W. Charge, orbital, and magnetic ordering in $\text{La}_{0.5}\text{Ca}_{0.5}\text{MnO}_3$. *Phys. Rev. B* **55**, 3015–3023 (1997).
- Su, Y., Du, C.-H., Hattori, P. D., Collins, S. P. & Cheong, S.-W. Charge ordering and the related structural phase transition in single-crystal $(\text{Bi}_{0.24}\text{Ca}_{0.76})\text{MnO}_3$. *Phys. Rev. B* **59**, 11687–11692 (1999).
- Liu, H. L., Cooper, S. L. & Cheong, S.-W. Optical study of the evolution of charge and spin ordering in the manganese perovskite $\text{Bi}_{1-x}\text{Ca}_x\text{MnO}_3$ ($x > 0.5$). *Phys. Rev. Lett.* **81**, 4684–4687 (1998).
- Billinge, S. J. L., DiFrancesco, R. G., Kwei, G. H., Neumeier, J. J. & Thomson, J. D. Direct observation of lattice polaron formation in the local structure of $\text{La}_{1-x}\text{Ca}_x\text{MnO}_3$. *Phys. Rev. Lett.* **77**, 715–718 (1996).
- Millis, A. J. Lattice effects in magnetoresistive perovskites. *Nature* **392**, 147–150 (1998).
- Woo, H., Tyson, T. A., Croft, M., Cheong, S.-W. & Woicik, J. C. Correlation between the magnetic and structural properties of Ca-doped BiMnO_3 . *Phys. Rev. B* **63**, 134412-1–134412-12 (2001).
- Radaelli, P. G., Cox, D. E., Capogna, L., Cheong, S.-W. & Marezio, M. Wigner-crystal and bi-stripe models for the magnetic and crystallographic superstructures of $\text{La}_{0.333}\text{Ca}_{0.667}\text{MnO}_3$. *Phys. Rev. B* **59**, 14440–14450 (1999).
- Matzdorf, R. *et al.* Ferromagnetism stabilized by lattice distortion at the surface of the p-wave superconductor Sr_2RuO_4 . *Science* **289**, 746–748 (2000).
- Kagan, M. Yu., Kugel, K. I. & Khomskii, D. I. Phase separation in systems with charge ordering. *J. Exp. Theor. Phys.* **93**, 415–423 (2001).

Acknowledgements

We thank N. Wingreen and J. Chadi for comments on the manuscript, and D. Khomskii, A. Millis, C. de Morais Smith and A. Yazdani for discussions. B.G.K. and S.W.C. were supported by the National Science Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to C.R. (e-mail: renner@research.nj.nec.com).

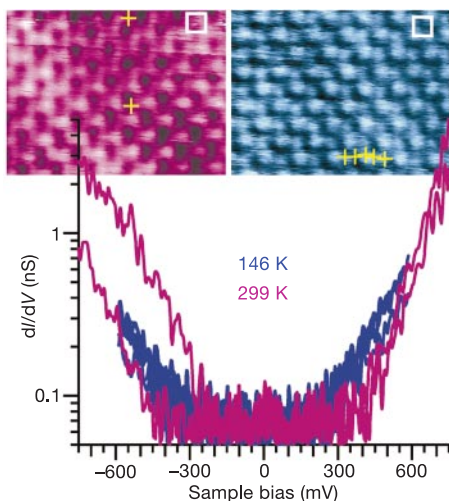


Figure 4 Identical lattice periods and electronic structures of the insulating $\sqrt{2}a_0 \times \sqrt{2}a_0$ regions observed in the paramagnetic room-temperature phase (left) and the charge-ordered low-temperature phase (right). The gap in the differential tunnelling conductance (numerical derivatives normalized to the 299 K junction resistance $R = V/I$ at 0.8 V) is essentially the same (~ 0.7 eV) at 299 K (purple) and at 146 K (blue). The STM constant-current images ($4.5 \times 3.5 \text{ nm}^2$) reveal the same lattice periods in regions where the insulating spectra were measured (yellow crosses). The white squares correspond to the cubic unit cell. For the 8 (out of 34) tungsten tips for which atomic resolution was obtained, there was 100% correspondence between the different structural phases and the conducting or insulating nature of the spectra.

Abiogenic formation of alkanes in the Earth's crust as a minor source for global hydrocarbon reservoirs

B. Sherwood Lollar, T. D. Westgate, J. A. Ward, G. F. Slater & G. Lacrampe-Couloume

Stable Isotope Laboratory, University of Toronto, Toronto, Ontario, Canada M5S 3B1

Natural hydrocarbons are largely formed by the thermal decomposition of organic matter (thermogenesis) or by microbial processes (bacteriogenesis). But the discovery of methane at an East Pacific Rise hydrothermal vent¹ and in other crustal fluids supports the occurrence of an abiogenic source of hydrocarbons^{2–4}. These abiogenic hydrocarbons are generally formed by the reduction of carbon dioxide, a process which is thought to occur during magma cooling⁵ and—more commonly—in hydrothermal systems during water–rock interactions, for example involving Fischer–Tropsch reactions and the serpentinization of ultramafic rocks^{6–10}. Suggestions that abiogenic hydrocarbons make a significant contribution to economic hydrocarbon reservoirs² have been difficult to resolve, in part owing to uncertainty in the carbon isotopic signatures for abiogenic versus thermogenic hydrocarbons^{4,10}. Here, using carbon and hydrogen isotope analyses of abiogenic methane and higher hydrocarbons in crystalline rocks of the Canadian shield, we show a clear distinction between abiogenic and thermogenic hydrocarbons. The progressive isotopic trends for the series of C₁–C₄ alkanes indicate that hydrocarbon formation occurs by way of polymerization of methane precursors. Given that these trends are not observed in the isotopic signatures of economic gas reservoirs, we can now rule out the presence of a globally significant abiogenic source of hydrocarbons.

Large volumes of methane gas discharge from fractures and exploration boreholes in hard rock mines operating throughout the Canadian and Fennoscandian shields. An abiogenic origin through water–rock interaction was proposed for these gases on the basis of $\delta^{13}\text{C}$ and $\delta^2\text{H}$ signatures for methane incompatible with bacterial or thermogenic methane¹¹. Recent experimental studies^{10,12,13} confirm that production of abiogenic methane by water–rock interaction can result in $\delta^{13}\text{C}$ values as depleted as those reported for methane from the above shields (–22.4‰ to –57.5‰)¹¹. Nonetheless, the evidence for an abiogenic origin for these shield gases was hitherto largely circumstan-

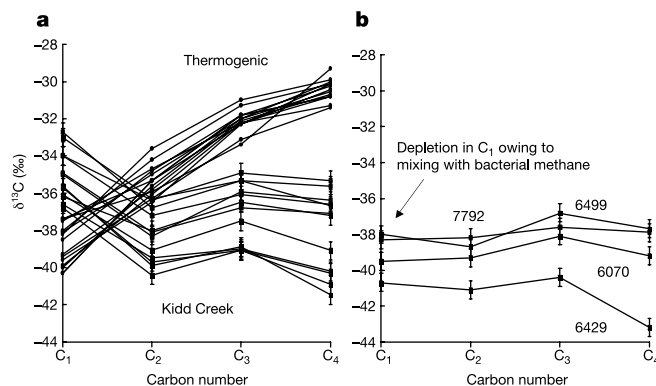


Figure 1 Plot of $\delta^{13}\text{C}$ values of individual *n*-alkanes against carbon number for gas samples from Kidd Creek mine, and for thermogenic gases from southwest Ontario natural-gas fields²⁹. Error bars on all $\delta^{13}\text{C}$ values are $\pm 0.5\text{‰}$. **a**, Of 17 boreholes, 13 show a trend of $\delta^{13}\text{C}$ depletion in C₂–C₄ with respect to C₁ (see text). **b**, For 4 of the 17 boreholes sampled at Kidd Creek, although the isotopic distribution still does not fit a typical thermogenic pattern, $\delta^{13}\text{C}_{\text{C}_2}$ is within error of $\delta^{13}\text{C}_{\text{C}_1}$, rather than depleted with respect to C₁. This is probably due to mixing of a Kidd Creek gas similar to those plotted in **a** with a component of bacterial methane such as reported in refs 18 and 19. For 8 of the 17 boreholes sampled in this study, time series data were collected over a period of 8–19 months. The isotopic compositions of C₁–C₄ alkanes for 7 of the boreholes are constant over time³⁰. For the 6070 borehole however, a small increase in the C₁/(C₂ + C₃) ratio and an isotopic depletion of 2.2‰ was observed for $\delta^{13}\text{C}_{\text{C}_1}$ over the 19-month period after its completion, consistent with addition of 10–25% bacterial methane (based on typical compositional and isotopic values for bacterial gases¹⁶). Such a phenomenon could account for the fact that for this sample, $\delta^{13}\text{C}_{\text{C}_1}$ is within error of $\delta^{13}\text{C}_{\text{C}_2}$ rather than significantly enriched with respect to C₂–C₄ as in the majority of the samples. No time series data are available for boreholes 6499, 6429 and 7792 as they were only sampled once, but they may reflect the same phenomenon. Certainly, whereas the 5 other boreholes from the 6,900-foot level have very similar $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values for C₁ (mean of 5 samples; $\delta^{13}\text{C} = -34.1 \pm 0.9\text{‰}$ and $\delta^2\text{H} = -413 \pm 5\text{‰}$), the isotopic values for C₁ from borehole 7792 are considerably more depleted in ^{13}C (–38.3‰) and enriched in ^2H (–390‰), consistent with mixing with a bacterial methane end-member (Table 1).

tial. In this study, the unusual pattern of $\delta^{13}\text{C}$ values among the C₁–C₄ alkanes provides evidence of abiogenic formation mechanisms (Table 1; Fig. 1a). Thermogenic hydrocarbons have been shown empirically and experimentally to have a characteristic isotope distribution pattern whereby the C₁–C₄ alkanes become more enriched in ^{13}C (less negative $\delta^{13}\text{C}$ values) with increasing molecular mass. This orderly isotopic distribution results from kinetic fractionation effects, whereby alkyl groups separating from source organic matter cleave preferentially at weaker ^{12}C – ^{12}C rather

Table 1 Carbon and hydrogen isotopic values for C₁–C₄ hydrocarbon gases

Borehole	$\delta^{13}\text{C}_{\text{C}_1}$	$\delta^2\text{H}_{\text{C}_1}$	$\delta^{13}\text{C}_{\text{C}_2}$	$\delta^2\text{H}_{\text{C}_2}$	$\delta^{13}\text{C}_{\text{C}_3}$	$\delta^2\text{H}_{\text{C}_3}$	$\delta^{13}\text{C}_{\text{C}_4}$	$\delta^2\text{H}_{\text{C}_4}$
6070	–39.5	NA	–39.3	NA	–38.1	NA	–39.2	NA
6080	–35.1	NA	–38.4	NA	–37.5	NA	–38.8	NA
6448	–33.0	NA	–36.3	NA	–35.3	NA	–36.7	NA
6499	–38.0	NA	–38.7	NA	–36.8	NA	–37.7	NA
6500A	–36.2	NA	–38.0	NA	–36.5	NA	–37.2	NA
6447	–35.6	NA	–39.9	NA	–38.9	NA	–41.5	NA
6500B	–36.1	NA	–39.1	NA	–37.5	NA	–39.1	NA
6298	–36.8	NA	–40.4	NA	–39.1	NA	–40.3	NA
6300	–35.7	NA	–39.7	NA	–39.1	NA	–40.9	NA
6429	–40.7	NA	–41.1	NA	–40.4	NA	–43.2	NA
6299	–36.6	NA	–39.5	NA	–39.0	NA	–40.2	NA
7792	–38.3	–390	–38.2	–299	–37.6	–256	–37.9	–224
8558	–32.7	–419	–36.8	–321	–35.3	–264	–35.7	–245
8428	–34.0	–412	–36.4	–314	–34.9	–262	–35.4	–236
8282	–35.0	–409	–38.3	–316	–35.9	–269	–36.3	–252
8402	–34.9	–406	–38.1	–319	–36.8	–267	–37.0	–234
8539	–34.0	–417	–37.2	–320	–36.1	–270	–36.8	–256

All values in ‰. NA, not analysed (see Methods).

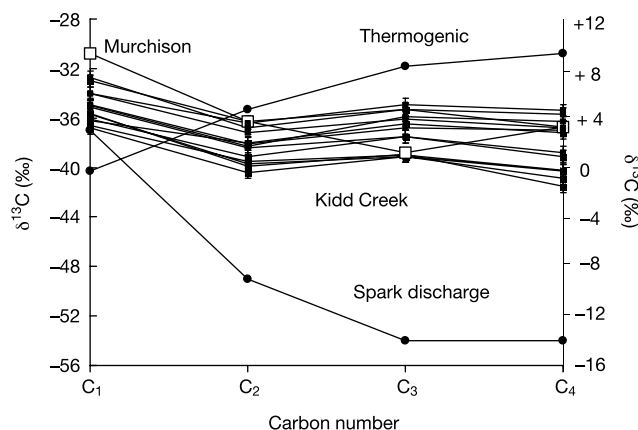


Figure 2 Plot of $\delta^{13}\text{C}$ values of individual n -alkanes against carbon number. $\delta^{13}\text{C}_{\text{C}_4}$ values include both iso- and n -butane. Error bars on $\delta^{13}\text{C}$ values for Kidd Creek samples are $\pm 0.5\text{‰}$. The isotopic pattern of C_1 – C_4 n -alkanes for the Kidd Creek samples (filled squares, left y-axis) are compared to the isotopic pattern for n -alkanes from the Murchison meteorite (open squares, right y-axis). $\delta^{13}\text{C}$ values (circles, left y-axis) for spark discharge experiments and thermogenic gas are from refs 14 and 29, respectively.

than ^{12}C – ^{13}C bonds¹⁴. This isotopic distribution pattern is ubiquitous among thermogenically derived gases, and is considered to be diagnostic of gases produced by the thermal decomposition of high-molecular-mass organic matter.

In contrast, Fig. 1a shows a significant depletion in ^{13}C for C_2 – C_4 with respect to C_1 for hydrocarbon gases from the Kidd Creek mine on the Canadian shield. For 13 of the 17 samples in this study, C_2 is significantly depleted in ^{13}C (by 2–4‰) with respect to C_1 (Fig. 1a). C_3 and C_4 are also isotopically depleted in ^{13}C compared to C_1 by 1–3‰ and by 1–6‰, respectively. No post-genetic alteration of typical thermogenic gas is known that could produce an isotopic enrichment in C_1 of the order of 10‰, and a simultaneous isotopic depletion in C_3 and C_4 relative to C_2 , to produce the pattern exhibited by the Kidd Creek gases^{15–17}. Although methanogenic bacteria have been identified in the ground waters of the Kidd Creek mine^{18,19}, addition of an isotopically depleted bacterial methane to the borehole gases would only reduce the C_1 enrichment observed in Fig. 1a and cannot account for the observed isotopic enrichment of C_1 relative to the higher-molecular-mass alkanes (Fig. 1b). A bacterial origin overall for the C_1 – C_4 gases is ruled out on the basis of C_1/C_2 ratios between 5.10 and 11.51, 2–4 orders of magnitude lower than for typical bacterial gas¹⁶.

Natural occurrences of hydrocarbon gas exhibiting isotopic depletion in C_2 – C_4 compared to C_1 are rare. Reported occurrences in the field include: C_1 – C_3 hydrocarbons in two gas samples from fluid inclusions associated with the Khibiny massif on the Kola peninsula, Russia²⁰; and C_1 – C_2 hydrocarbons from six gas samples from fluid inclusions associated with the Ilimaussaq complex, South Greenland^{21,22}. Unfortunately, none of these studies were able to obtain a significant number of $\delta^{13}\text{C}$ values for homologues higher than C_2 . As apparent isotopic depletion of C_2 versus C_1 can occur owing to bacterial oxidation of methane, conclusive evidence of isotopic depletion of higher hydrocarbons with respect to C_1 requires comparison to a larger range of higher hydrocarbons, particularly C_3 and C_4 . Only one isotopic data set exists for C_1 – C_4 alkanes of undisputed abiogenic origin. Values of $\delta^{13}\text{C}$ obtained for C_2 – C_4 n -alkanes for the Murchison meteorite show an isotopic depletion of 5–8‰ with respect to C_1 (Fig. 2). The absolute $\delta^{13}\text{C}$ values of the hydrocarbons from this meteorite are significantly more enriched in ^{13}C than the Kidd Creek samples, reflecting their extraterrestrial origin. Nonetheless, the Murchison and Kidd Creek samples show remarkable similarities in the isotopic distribution pattern between the C_1 – C_4 n -alkanes (Fig. 2). Although C_3 in the

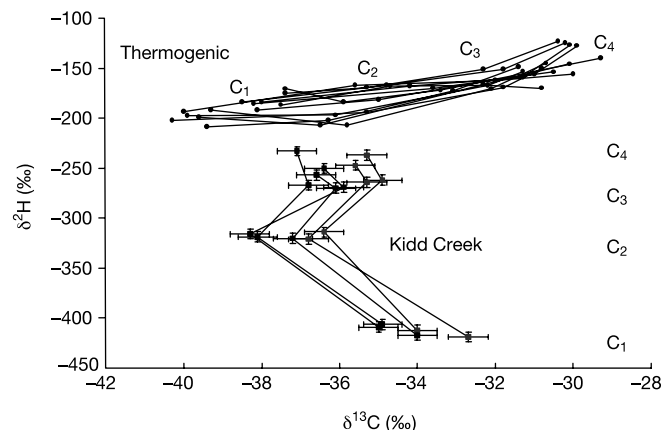


Figure 3 Plot of $\delta^{13}\text{C}$ versus $\delta^2\text{H}$ values for C_1 – C_4 for the Kidd Creek samples, and for thermogenic gases from southwest Ontario natural-gas fields²⁹. Error bars on Kidd Creek $\delta^{13}\text{C}$ values are $\pm 0.5\text{‰}$. Error bars on $\delta^2\text{H}$ values are $\pm 5\text{‰}$.

Murchison sample is more depleted relative to C_2 and C_4 than for the Kidd Creek samples, both sample sets show (1) significant depletion in $\delta^{13}\text{C}$ values of C_2 – C_4 versus C_1 , and (2) an isotopic pattern for C_1 – C_4 alkanes distinctly different from the typical pattern for thermogenic gases.

Such depletion patterns in C_2 – C_4 alkanes relative to C_1 have been shown to be produced by polymerization reactions, such as production of higher hydrocarbons from methane in a spark discharge experiment¹⁴ (Fig. 2). This pattern results during kinetically controlled synthesis of higher-molecular-mass homologues from lower ones owing to the fact that $^{12}\text{CH}_4$ reacts faster than $^{13}\text{CH}_4$ to form chains, so that ^{12}C is more likely to be incorporated into larger hydrocarbon chains^{14,23}. In other abiogenic polymerization reactions such as the Fischer–Tropsch synthesis, C_2 – C_4 hydrocarbons produced from CO and H_2 are also isotopically depleted with respect to C_1 , but the pattern of isotopic reversal ($\delta^{13}\text{C}_{\text{C}_1} > \delta^{13}\text{C}_{\text{C}_2} > \delta^{13}\text{C}_{\text{C}_3} = \delta^{13}\text{C}_{\text{C}_4}$) is not always as consistent as for the spark discharge experiment^{12,13}. Experimental results to date certainly show that whether by spark discharge or Fischer–Tropsch synthesis, hydrocarbons produced by abiogenic polymerization reactions do not show the isotopic enrichment trends for C_1 – C_4 typical of thermogenic gases. So although the $\delta^{13}\text{C}$ pattern between C_3 and C_4 in the Murchison sample is not identical to the spark discharge experimental results, a similar abiogenic polymerization reaction has been invoked²³ to explain the overall C_2 – C_4 isotopic depletion pattern for the Murchison meteorite n -alkanes. For the Kidd Creek samples, the pronounced isotopic depletion of C_2 – C_4 , and the lack of a thermogenic-type enrichment trend for C_1 – C_4 are as consistent with kinetically controlled polymerization reactions as the Murchison n -alkanes. Potential mechanisms for abiogenic gas synthesis in this geologic environment include: surface-catalysed polymerization from reduction of CO in the Fischer–Tropsch synthesis²⁴; heating or metamorphism of graphite–carbonate-bearing rocks^{25,26}; or other vapour–water–rock alteration reactions in the presence of catalytically active metals²⁷. Confirmation of the exact mechanisms in natural systems will be dependent on constraining possible reactions with additional experimental data. Regardless of the specific mechanism, the Kidd Creek gases are (to our knowledge) the first reported terrestrial samples to demonstrate this consistent depletion of C_2 – C_4 with respect to C_1 indicative of abiogenic formation reactions.

The $\delta^2\text{H}$ values for the Kidd Creek hydrocarbon gases provide additional support for an abiogenic origin. Figure 3 illustrates the positive correlation of $\delta^{13}\text{C}$ and $\delta^2\text{H}$ values that is an established pattern for thermogenic gas reservoirs worldwide, due to increasing isotopic enrichment in both ^{13}C and ^2H with increasing molecular

mass for the C_1 – C_4 homologues¹⁶. In contrast, both the isotopically depleted absolute δ^2H values, and the relationship between $\delta^{13}C$ and δ^2H values for the Kidd Creek C_1 – C_4 alkanes, are distinctly different from those of thermogenic gas. In particular, the inverse correlation of $\delta^{13}C$ and δ^2H values between C_1 and C_2 supports the participation of these compounds in an abiogenic polymerization reaction. It has been proposed¹⁴ that in a kinetically controlled synthesis of higher-molecular-mass homologues from lower ones, the lighter isotope (^{12}C) will react faster than the heavy isotope (^{13}C) to form a 2-carbon chain. ^{12}C would be more likely to be incorporated into the product of the carbon addition reaction, and the resulting ethane would be depleted in ^{13}C versus the methane precursor. By analogy, owing to preferential cleavage of the weaker ^{12}C – 1H bond versus the ^{13}C – 1H bond, the light (1H) isotope would be preferentially eliminated in a polymerization reaction. Hence, the resulting ethane would be expected to be isotopically enriched in 2H , and isotopically depleted in ^{13}C , with respect to the methane precursor. This inverse relationship of ^{13}C isotope depletion and 2H isotope enrichment between C_1 and C_2 for the Kidd Creek samples (Fig. 3) supports a polymerization reaction as the first step in the formation of the Kidd Creek higher hydrocarbons from a methane precursor.

The distinct carbon isotopic depletion trends observed for the Kidd Creek gases, and their similarity to the trends observed for the Murchison meteorite n -alkanes, provides evidence for production of these Canadian shield gases by abiogenic synthesis rather than by conventional thermogenic or bacteriogenic processes (Fig. 2). The pattern of ^{13}C isotopic depletion and 2H enrichment between C_1 and C_2 in the Kidd Creek gases (Fig. 3) fits a model of isotopic fractionation involving hydrocarbon formation by kinetically controlled polymerization reactions. This study demonstrates that abiogenic processes that may have generated prebiotic organic molecules in the early Earth continued to play a significant role in the production of hydrocarbons in the crystalline rocks of the Canadian shield. The contrasting range of δ^2H values, and the markedly different relationship of $\delta^{13}C$ and δ^2H values demonstrated for abiogenic C_1 – C_4 compared to thermogenic natural-gas reservoirs provide important criteria for distinguishing between these two sources. Based on the isotopic characteristics of abiogenic gases identified in this study, the ubiquitous positive correlation of $\delta^{13}C$ and δ^2H values for C_1 – C_4 hydrocarbons in economic gas reservoirs world-wide is not consistent with any significant contribution from abiogenic gas. □

Methods

Kidd Creek mine, situated in the southern volcanic zone of the Abitibi greenstone belt (approximately 2,700 Myr old), 24 km north of Timmins, Ontario, is one of the world's largest volcanogenic massive sulphide deposits. All gas samples were collected from the 6,800- and 6,900-foot levels (2,072–2,100 m below land surface), from boreholes drilled approximately perpendicular to the steeply dipping felsic, mafic and ultramafic units²⁸. Gases occur in association with saline ground waters and brines in pressurized 'pockets' formed by sealed fracture systems within the rocks. During exploration drilling, these systems are ruptured, triggering depressurization and gas release at rates as high as 30 l min⁻¹ per borehole. Samples were collected from freely discharging boreholes in evacuated glass sample vessels. Boreholes were sealed with packers, and samples collected from the packered interval to minimize any air contamination.

Gases are composed of methane (C_1 , 69.3–78.1%), ethane (C_2 , 5.55–11.7%), H_2 (0.40–12.7%) and N_2 (3.88–12.7%), along with minor concentrations of helium (1.83–2.45%), propane (C_3 , 0.71–2.42%) and butane (C_4 , 0.21–0.77%). CO_2 concentrations are all below detection. Compositional data are available as Supplementary Information. Compositional analyses were carried out as described in ref. 11. On the basis of crustal $^3He/^4He$ ratios for the shield gases, no mantle-derived component for the hydrocarbon gases is suggested¹¹.

Carbon isotopes were measured by compound specific isotope analysis using a Finnigan MAT 252 mass spectrometer interfaced with a Varian 3400 GC. Accuracy and reproducibility of $\delta^{13}C$ values are $\pm 0.5\%$ with respect to VPDB standard. $\delta^{13}C_{C4}$ values include both n -butane and isobutane for all samples. Hydrogen isotopes were measured by compound specific isotope analysis using a Finnigan MAT delta + XL mass spectrometer interfaced with an HP6890 gas chromatograph. Accuracy and reproducibility of δ^2H values are $\pm 5\%$ with respect to VSMOW standard. 6000 series boreholes were sampled from the 6,800-foot level (2,072 m below land surface) of Kidd Creek mine in 1996–97. As this predates the availability of continuous flow compound specific hydrogen isotope

analysis, only $\delta^{13}C$ values could be analysed. 7,000 and 8,000 series boreholes were sampled from the 6,900-foot level (2,100 m below land surface) of Kidd Creek mine in 2000. Both $\delta^{13}C$ and δ^2H values are reported for these samples, and δ^2H_{C4} values include both n -butane and iso-butane.

Received 22 November 2001; accepted 18 February 2002.

- Welhan, J. A. & Craig, H. Methane and hydrogen in East Pacific Rise hydrothermal fluids. *Geol. Res. Lett.* **6**, 829–831 (1979).
- Gold, T. Terrestrial sources of carbon and earthquake outgassing. *J. Petrol. Geol.* **1**, 1–19 (1979).
- Shock, E. L. An open or shut case? *Nature* **378**, 338–339 (1995).
- Jenden, P. D., Hilton, D. R., Kaplan, I. R. & Craig, H. In *The Future of Energy Gases* (ed. Howell, D.) 31–56 (Professional Paper 1570, US Geological Survey, 1993).
- Kelley, D. S. Methane-rich fluids in the oceanic crust. *J. Geophys. Res.* **101**, 2943–2962 (1996).
- Berndt, M. E., Allen, D. E. & Seyfried, W. E. J. Reduction of CO_2 during serpentinization of olivine at 300°C and 500 bar. *Geology* **24**, 351–354 (1996).
- Abrajano, T. A. *et al.* Geochemistry of reduced gas related to serpentinization of the Zambales ophiolite, Philippines. *Appl. Geochem.* **5**, 625–630 (1990).
- Vanko, D. A. & Stakes, D. S. Fluids in oceanic layer 3: evidence from veined rocks, hole 735B, Southwest Indian Ridge. *Proc. ODP Sci. Res.* **118**, 181–215 (1991).
- Charlou, J. L. & Donval, J. P. Hydrothermal methane venting between 12N and 26N along the Mid-Atlantic ridge. *J. Geophys. Res.* **98**, 9625–9642 (1993).
- Horita, J. & Berndt, M. E. Abiogenic methane formation and isotopic fractionation under hydrothermal conditions. *Science* **285**, 1055–1057 (1999).
- Sherwood Lollar, B. *et al.* Abiogenic methanogenesis in crystalline rocks. *Geochim. Cosmochim. Acta* **57**, 5087–5097 (1993).
- Yuen, G. U. *et al.* Carbon isotope fractionation in Fischer-Tropsch type reactions. *Lunar Planet. Sci. Conf. XXI*, 1367–1368 (1990).
- Hu, G., Ouyang, Z., Wang, X. & Wen, Q. Carbon isotopic fractionation in the process of Fischer-Tropsch reaction in primitive solar nebula. *Sci. China* **41**, 202–207 (1998).
- Des Marais, D. J., Donchin, J. H., Nehring, N. L. & Truesdell, A. H. Molecular carbon isotopic evidence for the origin of geothermal hydrocarbons. *Nature* **292**, 826–828 (1981).
- James, A. T. Correlation of natural gas by use of carbon isotopic distribution between hydrocarbon components. *Am. Assoc. Petrol. Geol. Bull.* **67**, 1176–1191 (1983).
- Schoell, M. Genetic characterization of natural gases. *Am. Assoc. Petrol. Geol. Bull.* **67**, 2225–2238 (1983).
- Pernaton, E., Prinzhofer, A. & Schneider, F. Reconsideration of methane isotope signature as a criterion for the genesis of natural gas. *Revue Inst. Français Petrole* **51**, 635–651 (1996).
- Doig, F. *Bacterial Methanogenesis in Canadian Shield Groundwaters*. Thesis, Univ. Toronto (1994).
- Doig, F., Sherwood Lollar, B. & Ferris, F. G. Microbial communities in deep Canadian Shield groundwaters—An in situ biofilm experiment. *Geomicrobiol. J.* **13**, 91–102 (1995).
- Khitrov, N. I. *et al.* Free emanation gases of the Khibiny massif. *Sov. Geol.* **2**, 62–73 (1979). (In Russian).
- Konnerup-Madsen, J., Kreulen, R. & Rose-Hansen, J. Stable isotope characteristics of hydrocarbon gases in the alkaline Ilimaussaq complex, south Greenland. *Bull. Mineral.* **106**, 642–653 (1988).
- Laier, T. & Nytoft, H. P. In *Proc. 17th Int. Meeting on Organic Geochemistry* (eds Grimalt, J. O. & Dorronsoro, C.) 1109–1111 (European Association of Organic Geochemists, San Sebastian, Spain, 1995).
- Yuen, G., Blair, N., DesMarais, D. J. & Chang, S. Carbon isotope composition of low molecular weight hydrocarbons and monocarboxylic acids from Murchison meteorite. *Nature* **307**, 252–254 (1984).
- Lancet, H. S. & Anders, E. Carbon isotope fractionation in the Fischer-Tropsch synthesis of methane. *Science* **170**, 980–982 (1970).
- Giardini, A. A. & Salotti, C. A. Synthesis of graphite and hydrocarbons by reaction between calcite and hydrogen. *Science* **159**, 317–319 (1968).
- Holloway, J. R. Graphite- CH_4 - H_2O - CO_2 equilibria at low grade metamorphic conditions. *Geology* **12**, 455–458 (1984).
- McCullum, T. M. & Seewald, J. S. A reassessment of the potential for reduction of dissolved CO_2 to hydrocarbons during serpentinization of olivine. *Geochim. Cosmochim. Acta* **65**, 3769–3778 (2001).
- Bleeker, W. & Parrish, R. R. Stratigraphy and U-Pb zircon geochronology of Kidd Creek: Implications for the formation of giant volcanogenic massive sulphate deposits and the tectonic history of the Abitibi greenstone belt. *Can. J. Earth Sci.* **3**, 1213–1231 (1996).
- Sherwood Lollar, B., Weise, S. M., Frap, S. K. & Barker, J. F. Isotopic constraints on the migration of hydrocarbon and helium gases of southwestern Ontario. *Bull. Can. Petrol. Geol.* **42**, 283–295 (1994).
- Westgate, T. D. *Compositional and Carbon Isotopic Characteristics of Abiogenically-derived Hydrocarbons at Kidd Creek Mine, Timmins, Ontario*. Thesis, Univ. Toronto (1998).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

This study was supported in part by Falconbridge Mining Ltd and by the Natural Sciences and Engineering Research Council of Canada. We thank N. Arner and N. VanStone, and the Geology Office at Kidd Creek mine (P. Olson, A. Coutts, R. Cook) for providing geological information and assistance with underground field work.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to B.S.L. (e-mail: bsl@quartz.geology.utoronto.ca).

Long-lived vortices as a mode of deep ventilation in the Greenland Sea

Jean-Claude Gascard*, Andrew J. Watson†, Marie-José Messias‡, K. Anders Olsson§, Truls Johannessen§ & Knud Simonsen||¶

* Université Pierre et Marie Curie, Laboratoire d'Océanographie Dynamique et de Climatologie, 75252 Paris cedex 05, France

† School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK

‡ Department of Analytical and Marine Chemistry, Göteborg University, S-412 96, Göteborg, Sweden

§ Geophysical Institute and the Bjerknes Centre for Climate Research, University of Bergen, Allegåten 55, N-5007 Bergen, Norway

|| Nansen Environmental and Remote Sensing Center, Edvard Griegsvei 3A, N-5037 Solheimsvik, Bergen, Norway

The Greenland Sea is one of a few sites in the world ocean where convection to great depths occurs^{1–4}—a process that forms some of the densest waters in the ocean. But the role of deep convective eddies, which result from surface cooling and mixing across density surfaces followed by geostrophic adjustment⁵, has not been fully taken into account in the description of the initiation and growth of convection⁶. Here we present tracer, float and hydrographic observations of long-lived (~1 year) and compact (~5 km core diameter) vortices that reach down to depths of 2 km. The eddies form in winter, near the rim of the Greenland Sea central gyre, and rotate clockwise with periods of a few days. The cores of the observed eddies are constituted from a mixture of modified Atlantic water that is warm and salty with polar water that is cold and fresh. We infer that these submesoscale coherent eddies contribute substantially to the input of Atlantic and polar waters to depths greater than 500 m in the central Greenland Sea.

As part of the European Sub-Polar Oceans programme (ESOP-2), an extensive series of oceanographic observations were made during the period 1996–97 in the Greenland Sea. Figure 1a shows the study area. During summer 1996, a release of sulphur hexafluoride (SF₆) tracer was made at ~300 m depth in the centre of the Greenland Sea⁷, and its fate followed by surveys in November 1996, and March and May 1997. Early in November 1996, four pairs of floats (one of each pair at 250 m and one at 500 m depth) were deployed along 74° 30' N at 2° W, 3° W, 4° W and 5° W respectively. Four more pairs of floats were deployed four months later throughout the gyre. Figure 1b shows the tracks of five of the 16 floats during the six months March–August 1997. All these floats show evidence of becoming entrained for most of their trajectories in anticyclonic eddies. The eddies tend to populate the rim of a region ~200 km across, the central gyre of the Greenland Sea, which circulates cyclonically and in which deep convection has been previously reported to occur^{1–4}.

Figure 2b shows the north–south component of the velocity of one of these floats. Float no. 02 was launched in November 1996. Late in December 1996 (day 50 of the experiment), it was entrained in an anticyclonic eddy at a time when sea-to-air heat fluxes increased drastically (Fig. 2a). On January 20 (day 80), this float moved in towards the eddy centre, spending 150 days circling the eddy centre at a constant temperature of –1 °C (Fig. 2d) within a radius of 2–3 km, before spiralling out again until its recovery in August 1997, indicating an eddy lifetime of at least nine months. A plot of the orbital speeds of the float against distance from the centre shows that the vortex had a 'core' of about 5 km diameter in which it rotated like a solid body, and a 'skirt' surrounding this out to radii of ~15 km in which angular velocity decreased. The core had relative

vorticity ξ equal to about $-f/2$, compared to $\xi = -f/8$ at 8 km radius in the skirt, where f is the planetary vorticity.

In May 1997 while the floats were circulating the eddy, hydrographic and tracer measurements of the same feature were made. Figure 3a and b show sections of density, and SF₆ tracer, passing through the eddy at 0° E along 75° N. The core of the eddy was composed of homogeneous water at about –1 °C (in agreement with float no. 02 temperatures, Fig. 2d), with high levels of chlorofluorocarbons (CFCs) and oxygen (O₂), and low SF₆ extending from approximately 300 m to 2,200 m depth.

Calculation of geostrophic currents confirms a strong anticyclonic spin to the core at mid-depth. Comparison with the float velocities shows that the geostrophic currents contributed about two-thirds of the total rotation, the discrepancy being accounted for by centrifugal force, the eddy being in 'gradient wind' balance. The eddy core was characterized by a Rossby number $R = 0.5$ ($R = \xi/f$), Burger number $B = 0.25$ ($B = (N^2/f^2)(H^2/L^2)$, where N is the Brunt-Väisälä frequency) and aspect ratio $\alpha = 0.25$ ($\alpha = H/L$, where H and L represent vertical and horizontal eddy dimensions respectively).

The observations of Fig. 2 indicate that the eddy resulted from an intrusion of –1 °C homogeneous water into a stratified rotating

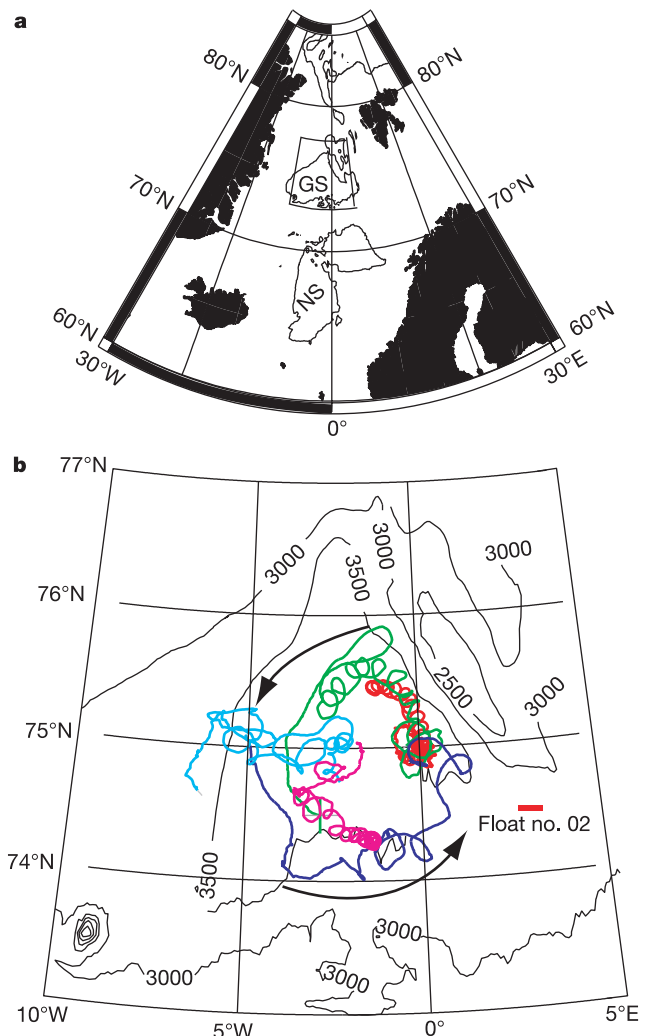


Figure 1 The Nordic seas (NS) and an enlargement of the central Greenland Sea (GS). **a**, The study area showing the 3,000 m depth contour, and **b**, six-month trajectories of five floats drifting at depths from 240 m to 530 m from March until August 1997 in the region marked by the box in **a**. Depth contours >2,500 m are also indicated.

¶ Present address: University of the Faroe Islands, PB 2109, FO-165 Argir, Faroe Islands.

environment, and that this intrusion occurred in January 1997 when air–sea heat fluxes were at their most intense. Models⁸ and rotating tank experiments⁹ have shown that intrusions of fixed volumes of homogeneous fluid into a stratified rotating medium produce anticyclonic lenses. These have interior cores that rotate as solid bodies, and approximately obey the relation $R \propto \sqrt{B}$ for a range of values of R from 0.1 to 0.4. This relationship fits the Greenland Sea eddy for which R varies from 0.5 near the centre, where maximum negative vorticity (very low potential vorticity) is observed, to 0.125 in the skirt. Hedstrom and Armi⁹ observed that lenses had a ‘fast spin-down’ phase in rotating tanks, and the ~70 rotations that float no. 02 executed in the eddy core correspond roughly to the expected lifetime of this phase. This may suggest that the float was ejected from the core as it began to collapse.

To locate the source of the water in the eddy core, we searched for binary combinations of surface and subsurface waters that would reproduce its salinity, temperature, and concentrations of CFCs, SF₆ and O₂ (Table 1). The high concentrations of oxygen and CFCs in the core of the eddy make it apparent that the water there contained a substantial component that had recently contacted the atmosphere. Scaling from these tracers, it is clear that the low SF₆ content in the eddy (~1 fmol l⁻¹) was also entirely of atmospheric origin, and not from the tracer release, despite the fact that the eddy was embedded in water containing relatively high concentrations (~40 fmol l⁻¹) of SF₆ from the release experiment. A combination

of ~36% surface Arctic waters (parent 2: P2 in Table 1) as encountered in March and May 1997 and ~64% ‘return Atlantic water’—relatively salty water of Atlantic origin recirculating southward along the East Greenland current (parent 1: P1)—which we encountered at ~500 m depth just west of the gyre in November 1996, fits the composition of the eddy core closely (Table 1). Regarding surface water, we have measurements for tracers in May 1997 but not for March 1997, which would have been preferable because it is closer to the time during which convection occurred (in January 1997). However, surface water in May 1997 should be a good representative for winter surface conditions, except that it had warmed by ~0.3 °C. In agreement with previous findings¹⁰, our observations indicate that surface salinity close to 34.80 p.s.u. at nearly freezing temperatures (~1.8 °C) is necessary to trigger mixing with subsurface saltier and warmer modified Atlantic waters (~34.90 p.s.u.). This produces a new vintage of Greenland Arctic Intermediate Water (GAIW), the densest water mass found within the Greenland Sea gyre in this salinity range.

Our observations suggest a pattern for deep convection substantially different from that described in ref. 6, in which eddies are not considered important in initiating or sustaining convection, but only in the dissipation of larger-scale ‘mixed patches’ formed by convection. Following air–sea heat loss, it appears that Greenland Sea submesoscale eddies are directly generated by geostrophic adjustment and diapycnal mixing (mixing across density surfaces) between surface polar waters and subsurface modified Atlantic waters. This results in intrusions of homogeneous GAIW characterized by very low potential vorticity in a stratified environment. A similar process has been proposed⁵ to be the most probable for generating anticyclonic submesoscale coherent vortices in the Labrador Sea. In our case, it would be expected to occur dominantly

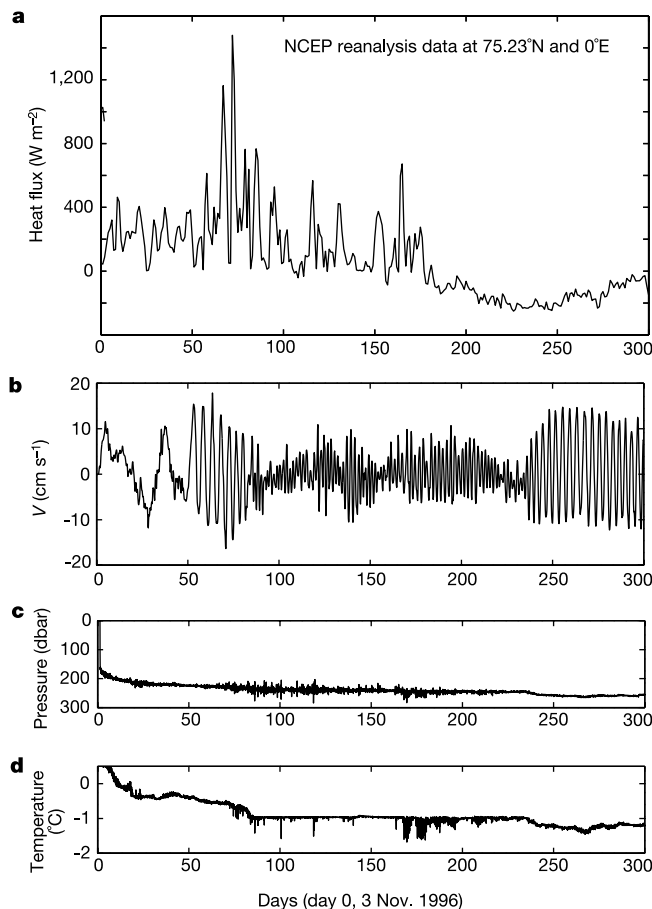


Figure 2 Time series for the period November 1996–August 1997. **a**, Total sea-to-air heat flux from NCEP (National Centers for Environmental Prediction) re-analysis; **b**, float no. 2 north-south horizontal velocity component; **c**, float no. 2 *in situ* pressure; **d**, float no. 2 *in situ* temperature. The reanalyses data were provided through the NOAA Climate Diagnostic Center at <http://www.cdc.noaa.gov>.

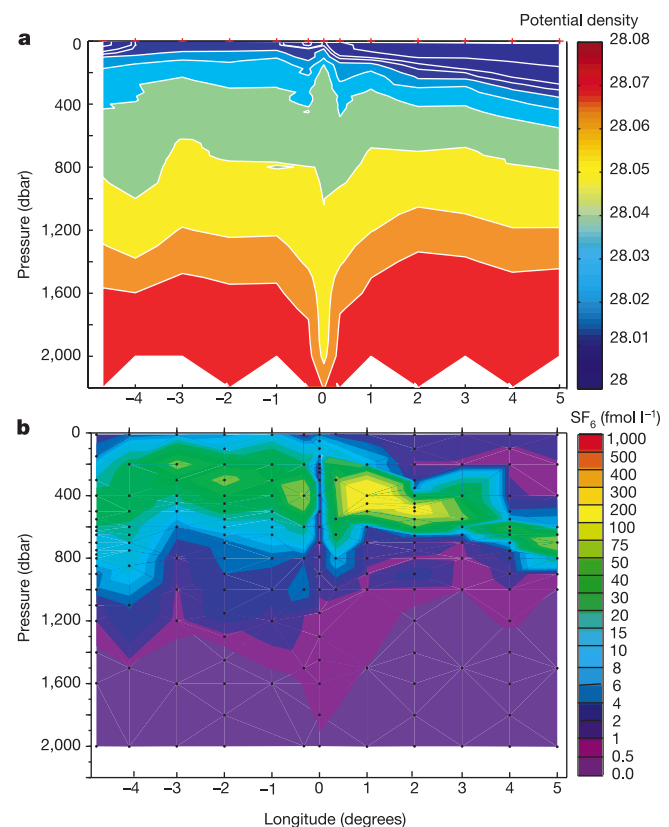


Figure 3 Vertical sections taken along 75°N in May 1997 in the central Greenland Sea. **a**, Seawater potential density (σ_t), and **b**, SF₆ concentration (fmol l⁻¹).

Table 1 Properties of the eddy core and its parent waters

	Return Atlantic water*, P1	Surface water†,‡ P2	Mixture, 64% P1 + 36% P2	Eddy core‡
Salinity (p.s.u.)	34.894	34.810	34.864	34.872
Oxygen ($\mu\text{mol kg}^{-1}$)	320.5	358.7	334.2	329.7
CFC-11 (pmol kg^{-1})	4.63	7.00	5.48	5.74
CFC-12 (pmol kg^{-1})	2.22	3.34	2.62	2.69
SF ₆ (fmol l^{-1})	0.78	1.61	1.08	1.17
Potential temperature (°C)	-0.47	-1.55† - 1.8§	-0.86 - 0.95	-0.998

Water mass characteristics observed at various locations and times.

* 74° 15' N, 6° W, 500 m depth, November 1996.

† 75° N, 2° W, 10 m depth, May 1997.

‡ 75° N, 0° E, average of 300–900 m depth, May 1997.

§ Lowest temperature observed in March 1997 during winter cooling nearer the time of convection.

at the rim of the gyre, where the parent originating from return Atlantic water and that coming from inside-gyre surface waters, meet and mix. Because these coherent eddies are stable for ~1 year, they may also precondition water masses for convective activity in the following winter season. They could then form foci to concentrate further convection^{4,11,12} after erosion of the layer of less dense water that caps the core during the summer.

On the basis of float data only, we have direct evidence for about eight such eddies during ESOP-2. This probably underestimates the true number, because almost every float we released that strayed into the rim region of the Greenland gyre in 1997 became entrained in an anticyclonic eddy. When these eddies eventually decay, they presumably release their core water at mid-depths in the Greenland Sea, ventilating the intermediate water. Each eddy core has a volume ~50 km³. Comparing the volume of water in eight eddy cores to a total volume involved in convection of $(2-4) \times 10^{12}$ m³ during 1996–97, a figure we have previously calculated from analysis of the tracer release experiment⁷, suggests a contribution of 10–20% to the total amount of convection from the eddies. However, only a small proportion—less than 10% of the total amount of water involved in convection—penetrated to around 1,000 m or deeper, most being confined to the upper ~500 m. Thus the eddies made a significant contribution to the total volume of deep convection—and dominated the water injected to substantial depth—in the Greenland Sea in 1996–97. Greenland Arctic Intermediate Water is thought to contribute to the overflows leading into the North Atlantic deep water, and the eddies thus provide a pathway for ventilation of the deep North Atlantic. Long-lived submesoscale anticyclonic vortices have also been observed in the Labrador Sea^{13,14} (another site known for deep ocean convection), indicating that such vortices may be ubiquitous features of deep ocean convection.

Received 19 October 2001; accepted 1 February 2002.

- Rudels, B., Quadfasel, D., Friedrich, H. & Houssais, M. N. Greenland Sea convection in the winter of 1987–1988. *J. Geophys. Res.* **94**, 3223–3227 (1989).
- Schott, F., Visbeck, M. & Fischer, J. Observations of vertical currents and convection in the central Greenland Sea during the winter of 1988–1989. *J. Geophys. Res.* **98**, 14401–14421 (1993).
- Morawitz, W. M. L., Sutton, P. J., Worcester, P. F. & Cornuelle, B. D. Three-dimensional observations of a deep convective chimney in the Greenland Sea during winter 1988/1989. *J. Phys. Oceanogr.* **26**, 2316–2343 (1996).
- Lherminier, P., Gascard, J.-C. & Quadfasel, D. The Greenland Sea in winter 1993 and 1994: Preconditioning for deep convection. *Deep Sea Res. II* **46**, 1199–1235 (1999).
- McWilliams, J. C. Vortex generation through balanced adjustment. *J. Phys. Oceanogr.* **18**, 1178–1192 (1988).
- Marshall, J. & Schott, F. Open-ocean convection: observations, theory and models. *Rev. Geophys.* **37**, 1–64 (1999).
- Watson, A. J. *et al.* Mixing and convection in the Greenland sea from a tracer-release experiment. *Nature* **401**, 902–904 (1999).
- Gill, A. E. Homogeneous intrusions in a rotating stratified fluid. *J. Fluid Mech.* **103**, 275–295 (1981).
- Hedstrom, K. & Armi, L. An experimental study of homogeneous lenses in a stratified rotating fluid. *J. Fluid Mech.* **191**, 535–556 (1988).
- Roach, A. T., Aagaard, K. & Carsey, F. Coupled ice-ocean variability in the Greenland Sea. *Atmos. Ocean* **31**, 319–337 (1993).

- Legg, S., McWilliams, J. & Gao, J. Localisation of deep ocean convection by a mesoscale eddy. *J. Phys. Oceanogr.* **28**, 944–970 (1998).
- Straneo, F. & Kawase, M. Comparisons of localized convection due to localized forcing and to preconditioning. *J. Phys. Oceanogr.* **29**, 55–68 (1999).
- Gascard, J.-C. & Clarke, R. A. The formation of Labrador Sea Water. Part 2: mesoscale and smaller-scale processes. *J. Phys. Oceanogr.* **13**, 1779–1797 (1983).
- Lilly, J. M. & Rhines, P. B. Coherent eddies in the Labrador Sea observed from a mooring. *J. Phys. Oceanogr. (special issue)* (in the press).

Acknowledgements

We thank the staff and crew of RV *Håkon Mosby*, RV *Johan Hjort* and RRS *James Clark Ross* for their support. The main support for this work was from the programmes of the EU: European sub-polar ocean project (ESOP), Tracer and Circulation in the Nordic Seas Region (TRACTOR) and Monitoring the Atlantic Inflow toward the Arctic (MAIA). Additional support from the following national agencies was also important: IFRTF (France), NERC (UK). ESOP required the involvement of a large number of individuals. We thank E. Jansen, coordinator of ESOP 2 and F. Rey for leading the RV *Johan Hjort* cruise; E. Fogelqvist, T. Tanhua, D. Wallace, C. Rouault and A. Lourenço for their contributions to the tracers and floats programmes respectively. NCEP-NCAR re-analysis data were provided through the NOAA Climate Diagnostics Center.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.-C.G. (e-mail: gascard@lodyc.jussieu.fr).

Fin development in a cartilaginous fish and the origin of vertebrate limbs

Mikiko Tanaka*, Andrea Münsterberg*†, W. Gary Anderson‡, Alan R. Prescott§, Neil Hazon‡ & Cheryl Tickle*

* Division of Cell and Developmental Biology, § Division of Cell Biology and Immunology, Wellcome Trust Biocentre, University of Dundee, Dow Street, Dundee DD1 5EH, UK

‡ School of Biology, Division of Environmental and Evolutionary Biology, Gatty Marine Lab, University of St Andrews, St Andrews, Fife, KY16 8LB, UK

Recent fossil finds and experimental analysis of chick and mouse embryos highlighted the lateral fin fold theory, which suggests that two pairs of limbs in tetrapods evolved by subdivision of an elongated single fin¹. Here we examine fin development in embryos of the primitive cartilaginous fish, *Scyliorhinus canicula* (dogfish) using scanning electron microscopy and investigate expression of genes known to be involved in limb positioning, identity and patterning in higher vertebrates. Although we did not detect lateral fin folds in dogfish embryos, *Engrailed-1* expression suggests that the body is compartmentalized dorso-ventrally. Furthermore, specification of limb identity occurs through the *Tbx4* and *Tbx5* genes, as in higher vertebrates. In contrast, unlike higher vertebrates, we did not detect *Shh* transcripts in dogfish fin-buds, although *dHand* (a gene involved in establishing *Shh*) is expressed. In *S. canicula*, the main fin axis seems to lie parallel to the body axis. 'Freeing' fins from the body axis and establishing a separate 'limb' axis has been proposed to be a crucial step in evolution of tetrapod limbs^{2,3}. We suggest that *Shh* plays a critical role in this process.

The continuous fin fold theory was once considered to be "more an established fact than a theory"³ but was subsequently questioned because of inconsistencies in the fossil record and in the embryology of cartilaginous fish⁴. Recently discovered fossils of the earliest-

† Present address: School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, UK.

known Cambrian vertebrates, however, seem to have had lateral ribbon-shaped fins⁵, and recent analysis of muscle formation in dogfish has shown uniform budding from somites in both limb and inter-limb regions⁶. The embryos of some primitive cartilaginous fish have been reported to have lateral fin folds⁷.

We therefore re-examined *Scyliohinus canicula* dogfish embryos from pre-fin-bud to fin-bud stages using scanning electron microscopy. In pre-fin-bud embryos (stage 22; ref. 8), no signs of lateral ridges could be detected (Fig. 1a). By stage 24 (about 3 days later), small shelf-like buds (pectoral fin buds) project out of the body wall just ventral to somites 5–13 (Fig. 1b). These buds are rimmed distally by a raised structure, the apical fold, and have discrete edges both anteriorly and posteriorly with no sign of continuation of the 'shelf' posteriorly along the flank. A thickening which will develop into pelvic fin buds is detected ventral to somites 27–36 (arrows; Fig. 1b). In later-stage embryos (stage 27, about 10–12 days older), pelvic fin buds (opposite somites 27–36) are present in addition to pectoral fin buds. Both types of bud are semicircular flaps rimmed with well defined apical folds. In between the buds, the body wall is quite smooth (Fig. 1c). Thus, we did not detect lateral folds in *S. canicula* embryos.

In chick embryos, the body ectoderm is compartmentalized with respect to the dorso-ventral axis in both limb-forming and flank regions⁹. This compartmentalization serves to position limbs laterally, because the apical ectodermal ridge, the thickened epithelium that mediates limb-bud outgrowth, arises at the compartment boundary. *Engrailed-1* expression is restricted to the ventral compartment of the body ectoderm of both chick and mouse embryos at pre-limb-bud stages and then in the ventral ectoderm and the ventral apical ridge of limb buds^{9,10}. To gain insight into ectoderm

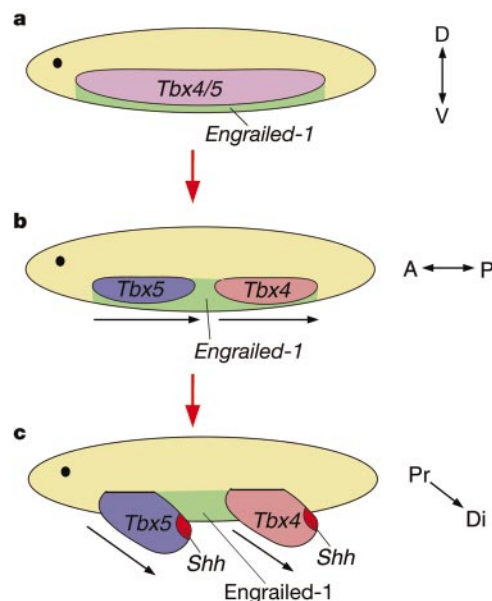


Figure 2 Model for evolution of vertebrate paired appendages. **a**, Hypothetical ancestral vertebrate with lateral fin folds, expressing *Tbx4/5*, positioned at the *Engrailed-1* expression boundary. **b**, *Tbx4/5* cluster duplication took place before the ancestor of cartilaginous fishes (or in that lineage) and inter-limb formation was suppressed. *Tbx5* and *Tbx4* are expressed in the pectoral and pelvic fins, respectively. **c**, Following the acquisition of *Shh* expression, the main axes of paired fins are freed from the body wall and establish the proximo-distal axes of the appendages. D, dorsal; V, ventral; A, anterior; P, posterior; Pr, proximal; Di, distal.

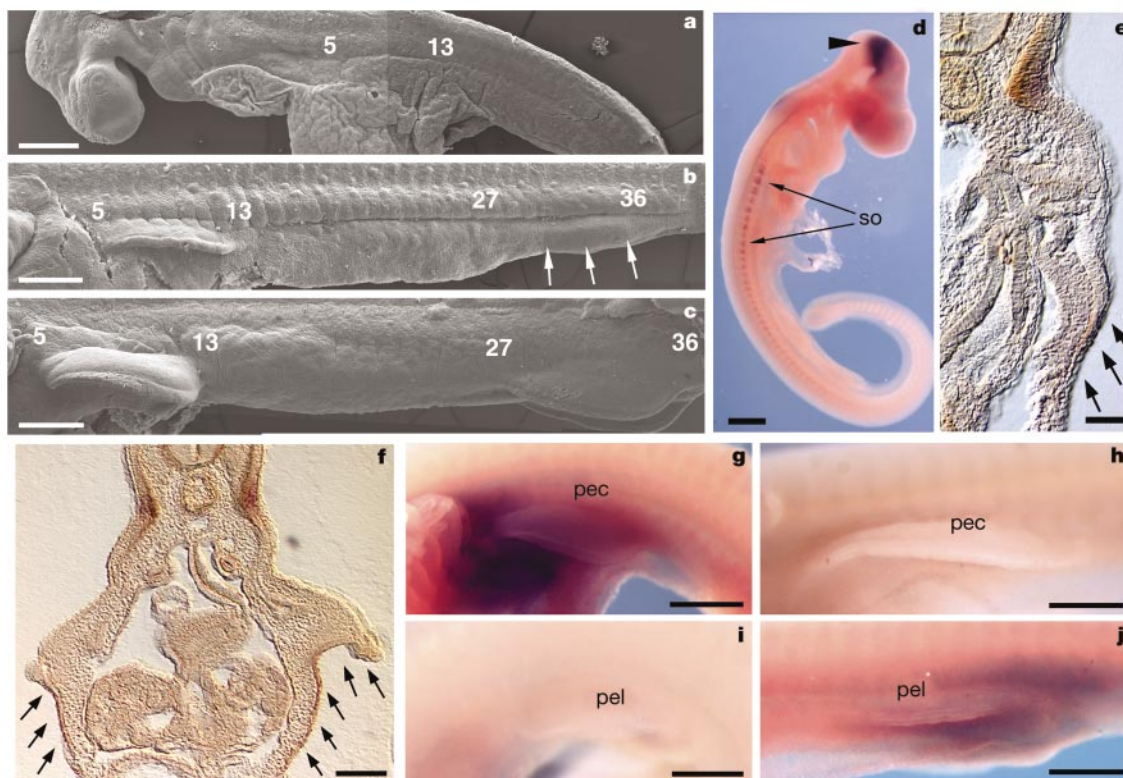


Figure 1 Scanning electron micrographs of *S. canicula* embryos. **a**, Stage 22; **b**, stage 24; and **c**, stage 27. **a–c**, Whole embryos. Pectoral and pelvic fins develop opposite somites 5–13 and 27–36 (arrows in **b**). Scale bars, 1 mm. **d**, Expression pattern of *ScEngrailed-1* in *S. canicula* embryos at stage 23. *ScEngrailed-1* expression is observed in the border of mesencephalon and metencephalon (arrowhead) and somites (arrows). Scale bar, 1 mm. **e**, **f**, Cross-sections of stage 23 and 26 embryos at pectoral fin-bud

level. *ScEngrailed-1* expression in presumptive fin buds region and fin buds indicated by arrows. Scale bars, 300 μ m in **e**, 500 μ m in **f**. **g–j**, *ScTbx4* and *ScTbx5* genes expression pattern in paired appendages in *S. canicula*. **g**, **i**, *ScTbx5* expression at stage 28. **g**, Pectoral fin. **i**, Pelvic fin. **h**, **j**, *ScTbx4* expression at stage 28. **h**, Pectoral fin. **j**, Pelvic fin. so, somite; pec, pectoral fin; pel, pelvic fin. Scale bars, 1 mm.

compartmentalization and whether fin buds of *S. canicula* arise at an ectodermal boundary of *Engrailed-1* expression, we isolated complementary DNA fragments of *ScEngrailed-1* and carried out whole-mount *in situ* hybridization on dogfish embryos. At both pre-fin (stage 23; Fig. 1d) and fin-bud stages, *ScEngrailed-1* transcripts were clearly seen at the border of the mesencephalon and metencephalon and in the somites. When these whole mounts were sectioned, *ScEngrailed-1* was seen to be expressed in the ventral ectoderm of the body of dogfish embryos both in the presumptive pectoral fin-bud region (stage 23 embryo; Fig. 1e) and in the pectoral fin buds themselves (stage 26 embryo; Fig. 1f). At the later stage, it was clear that, in the pectoral fin, the boundary between *ScEngrailed-1*-expressing ectoderm and non-*ScEngrailed-1*-expressing ectoderm lies at the mid-point of the apical fold, thus respecting the same boundary as *Engrailed-1* expression in chick and mouse limb buds. Thus, the pattern of *Engrailed-1* expression in *S. canicula* in relation to fin development suggests that the body of the dogfish is compartmentalized as in higher vertebrates. The mechanism that determines the dorso-ventral position of tetrapod limbs appears to be an ancient feature of the gnathostome body plan and could form the basis for a lateral fin fold in an ancestral vertebrate (Fig. 2a).

According to fin fold theory, two pairs of tetrapod limbs evolved by subdivision of a single lateral fin. In contrast, other recent ideas suggest that an ancestral vertebrate had a single pair of fins; either posterior¹¹ or anterior¹². In tetrapods and teleosts, the identity of each pair of appendages is determined by expression of T-box genes *Tbx5* and *Tbx4*, which are expressed in the anterior and posterior appendages, respectively¹³. In contrast, the cephalochordate *Amphioxus* has only one such gene, *AmphiTbx4/5* (ref. 14), although its expression pattern has not been reported. *Tbx4* and *Tbx5* genes in tetrapods are thought to have arisen by duplication of

this single ancestral gene. Therefore, we analysed the T-box gene complement in dogfish by employing polymerase chain reaction after reverse transcription of RNA (RT-PCR), using two sets of degenerate oligonucleotide primers. With each set of primers we obtained distinct PCR products, showing that the dogfish has both *Tbx4* and *Tbx5* genes. The deduced amino-acid sequences of *S. canicula Tbx4* and *Tbx5* T-box regions, aligned with those of zebrafish, chick, human homologues and the T-box region of *AmphiTbx4/5* are shown in Supplementary Information. The T-box region alignment provides further evidence that *ScTbx4* and *ScTbx5* belong to *Tbx4* and *Tbx5* subfamilies respectively. Skate embryos may also have a *Tbx5* gene¹³. In stage 28 dogfish embryos, *ScTbx5* expression was observed in pectoral fin buds (Fig. 1g), dorsal eye and heart, but not in pelvic fin buds (Fig. 1i). On the other hand, *ScTbx4* expression was not detected in pectoral fin buds (Fig. 1h), but was seen in pelvic fin buds (Fig. 1j). These results show that a *Tbx4/5* gene cluster duplication occurred before the origin of cartilaginous fish (Fig. 2b). Furthermore, as in tetrapods and teleosts, *Tbx5* and *Tbx4* are expressed in the anterior and posterior appendages, respectively.

An important recent finding with respect to the fin fold theory is that the inter-limb region in tetrapods has the potential to form limbs. Balinsky first discovered this in newts¹⁵ but limb-forming potential has now been described in both chick and mouse embryos^{16,17}. This remarkable property is associated with the potential of flank cells to form a polarizing region¹⁷. The polarizing region is the major limb 'organizer' and consists of mesenchyme at the posterior limb bud margin that expresses *Shh*¹⁸. We isolated an *ScShh* fragment using degenerate PCR and used this fragment to examine the *ScShh* expression pattern in stage 27 (Fig. 3a–c) and 28 (not shown) dogfish embryos. *ScShh* expression was observed in the floor plate, the notochord (Fig. 3b), the branchial arches (Fig. 3c),

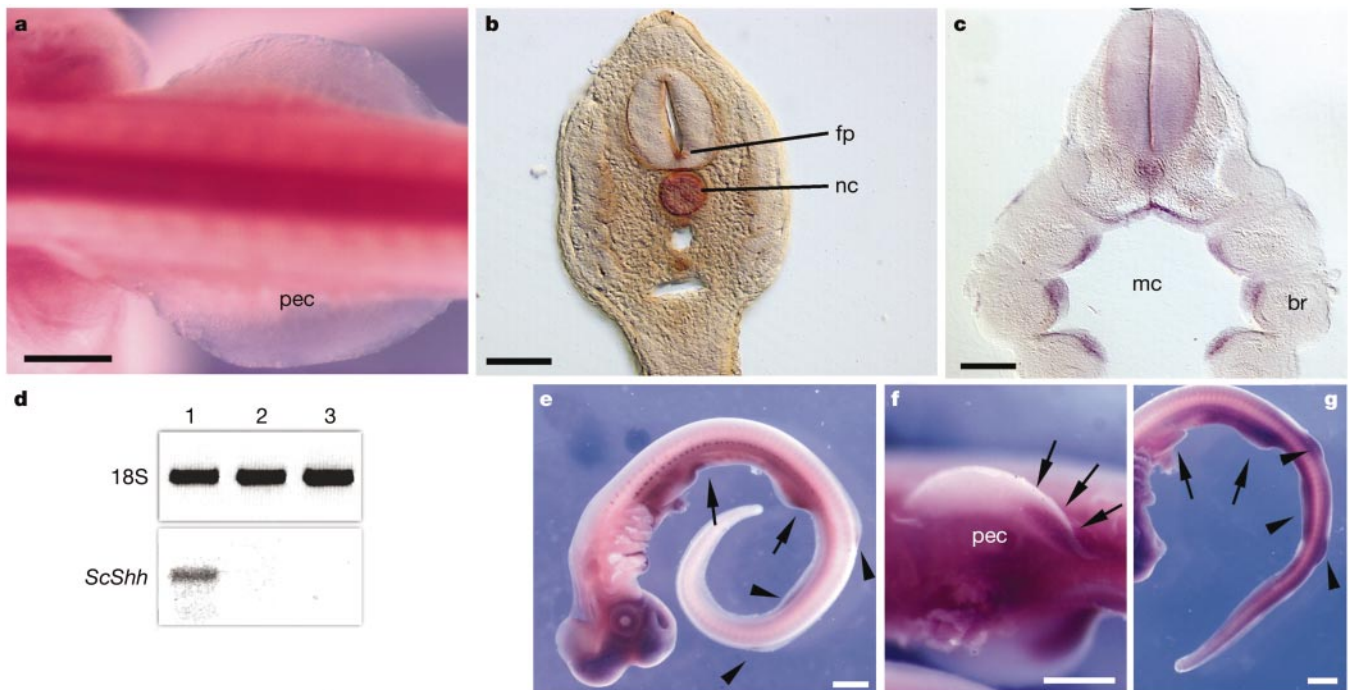


Figure 3 Gene expression patterns of stage 27–28 *S. canicula* embryos.

a–c, Expression pattern of *ScShh* in stage 27 *S. canicula* embryos. *ScShh* is expressed in the floor plate, the notochord and the branchial arches (**b** and **c**), but not in the pectoral fin bud (**a**). Scale bars, 500 μ m. **d**, *ScShh* expression of the stage 27 embryo was analysed by RT-PCR. *ScShh* expression was observed from body tissue (lane 1), but not from either pectoral (lane 2) nor pelvic fin-bud tissue (lane 3). 18S rRNA expression is shown as

an internal control. **e, f**, Expression pattern of *ScdHand* in stage 27 *S. canicula* embryo. *ScdHand* expression was observed in the posterior part of paired fin buds (arrows) and the dorsal and anal fin buds (arrowheads). Scale bars, 1 mm. **g**, Expression pattern of *ScBmp4* in stage 28 *S. canicula* embryo. *ScBmp4* expression was observed throughout paired fin buds (arrows) and dorsal and anal fin buds (arrowheads). pec, pectoral fin; fp, floor plate; nc, notochord; mc, mouth cavity; br, branchial arch. Scale bar, 1 mm.

the brain and the optic vesicle, but surprisingly we could not detect *Shh* expression in either pectoral (Fig. 3a) or pelvic fin buds. We extended the length of the *ScShh* probe using another set of primers, but even with this longer probe we were unable to detect a signal in fin buds. We further attempted to detect *ScShh* transcripts in stage 27 and 28 embryos by RT-PCR but although we were able to detect *ScShh* in the body, *ScShh* was not detected in either pectoral or pelvic fin buds (Fig. 3d). In zebrafish (*Danio rerio*), another member of the hedgehog gene family, *Tiggy-winkle hedgehog* (*Twhh*), has been reported¹⁹. It seems unlikely that the *Shh* we have isolated is *Twhh*

because *Twhh* is expressed only in the floorplate in zebrafish embryos¹⁹ and *ScShh* transcripts were observed in both the floorplate and the notochord in *S. canicula* embryos (Fig. 3b). Thus it appears that *Shh* must be expressed at very low levels, if at all, in the fin buds of dogfish embryos.

In higher vertebrates, genes that are involved in regulation of *Shh* expression (for example *dHand*²⁰) and that are targets of *Shh* signalling (for example, *Bmp4*; ref. 21) have been identified. We cloned these two genes from dogfish by degenerate PCR and examined their expression patterns. *ScdHand* transcripts were found in the same regions as in higher vertebrates, including the heart, the head and the posterior part of the fin buds where the polarizing region normally develops (Fig. 3e and f). In addition, *ScdHand* was detected at the posterior margins of the dorsal and anal fin buds (Fig. 3e). In higher vertebrates, *Bmp4* transcripts are abundant in anterior mesenchyme and expression is negatively regulated by *Shh*²¹. In dogfish fin buds, *ScBmp4* transcripts were detected throughout the fin bud (Fig. 3g), as might be predicted if *ScShh* expression is absent.

Our inability to detect *Shh* expression in dogfish fin buds was at first surprising. It could be that the acquisition of *Shh* expression in appendages was of primary importance for evolution of the distal regions. However, another possibility is that *Shh* could play a role in the 'freeing' of fins and the establishment of a proximo-distal limb axis (Fig. 2c). In *S. canicula*, as in many other cartilaginous fish, the metapterygium—the main long bone of the fin—develops parallel to the main body axis, in the proximal part of the fin bud next to the body wall (Fig. 4A). One hypothesis, which is still controversial, is that the metapterygial axis rotated outwards from the body wall during evolution^{2,3} (Fig. 4A). Fate maps of cells at the base of a chick wing-bud show that anterior cells remain at the proximal part of the limb, but cells in the middle and posterior regions extend posteriorly and distally as the wing bud grows out (Fig. 4B–D). The relative displacement of these populations is illustrated by the curved red line in Fig. 4C. Support for the idea that *Shh* is involved in this reorientation comes unexpectedly from a recent detailed analysis of *Shh* null mutant mouse embryos, which described the limbs as being 'trapped' within the body wall²².

Our data suggest that the dogfish may represent an important intermediate form in the evolution of tetrapod limbs (Fig. 2b). Although we did not detect lateral fin folds in dogfish, we suggest that an ancestral vertebrate had already been compartmentalized dorso-ventrally with *Engrailed-1* expression ventrally and had lateral fin folds that expressed a single *Tbx4/5* gene (Fig. 2a). This assumes, however, that a lateral fin fold is likely to be primitive. It will be interesting to characterize these features of the body plan in hagfish and lampreys. Our view of the origin of tetrapod limbs also suggests that more attention should be given to mechanisms that inhibit limb development. Many genes that are thought to be involved in limb formation in higher vertebrates, including *Fgf10* and *Hoxd9*, are initially expressed in both limb and inter-limb regions and then 'switch off' in the inter-limb regions. □

Methods

Scanning electron microscopy

S. canicula eggs were incubated at 16 °C in sea water and staged according to ref. 8. Embryos were fixed in Peter's fixative (1.25% glutaraldehyde, 1% paraformaldehyde in 0.1 M cacodylate buffer, pH 7.2) at 4 °C. After post-fixation in 1% osmium in cacodylate buffer, specimens were dehydrated in graded ethanol, placed in acetone, critically point-dried, sputter-coated with gold/palladium and viewed on a Hitachi S-4700 field-emission scanning electron microscope.

Identification of *S. canicula* gene homologues

We identified fragments of *S. canicula* *Tbx4* (275 base pairs, bp), *Tbx5* (278 bp), *Engrailed-1* (326 bp), *dHand* (267 bp), *Bmp4* (243 bp) and *Shh* (297 bp) from complementary DNA pools prepared from stage 24–28 embryos⁸ using degenerate primers. The amino-acid sequences used for degenerate primers were: *ScTbx4*, KFCNDKWM and TAFCTHVF; *ScTbx5*, YKFADNKW and TAFCTHVF; *ScEngrailed-1*, WPAWVYCT and MAQGLYNH; *ScdHand*, ECIPNVP and WPQHVA; *ScBmp4*, WIVAPP and DMVVEG. To avoid

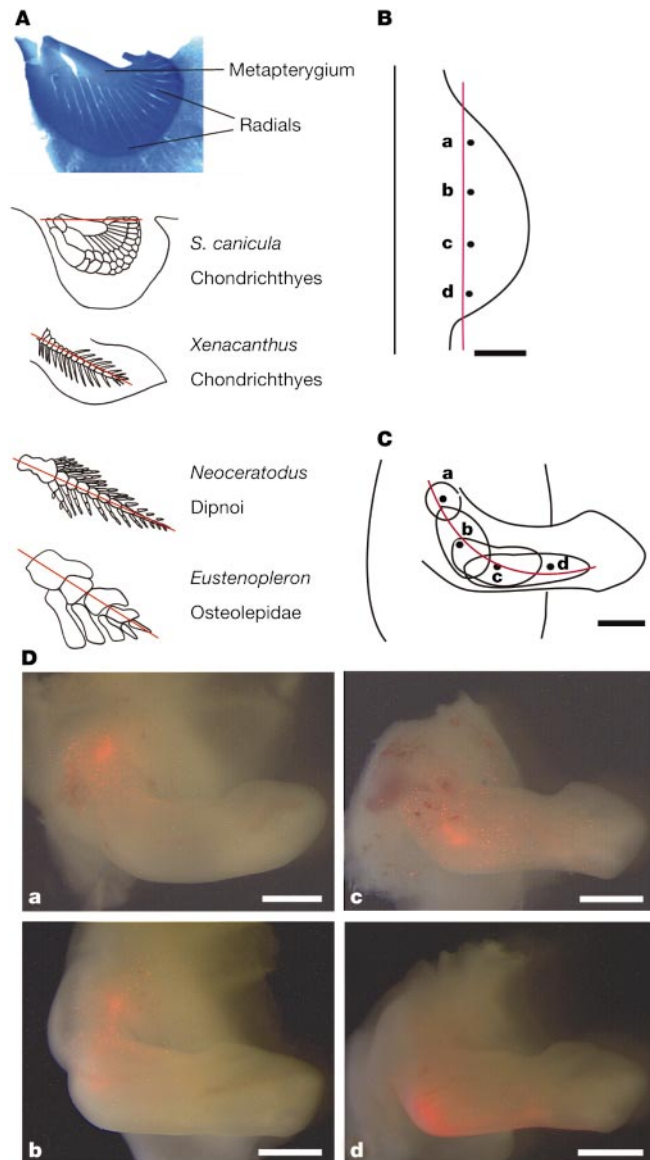


Figure 4 **A**, Series of pectoral appendages comparing the dogfish with other species in which the metapterygial axis is freed from the body wall and forms the main axis of the limb. The red line represents the position of the major axis. In *S. canicula*, the metapterygial stem has retained its original position in the body wall; in the fossils *Xenacanthus* and *Neoceratodus*, the metapterygial stem has become freed from the body wall and forms the principal axis^{3,28}. **B**, Diagram of stage 20 chick wing-bud showing positions at which Dil was injected. Scale bar, 250 μm. **C**, Map of representative results obtained in chick wing at 96 h for populations of labelled cells. Scale bar, 500 μm. The relative displacement of cells is illustrated by the red line. **D**, Micrographs showing fate maps of wing buds 96 h after Dil injection at the positions a–d respectively indicated in **B** and **C**. Scale bars 500 μm.

amplification of *Dhh* and *Ihh*, the *ScShh* fragment was obtained by nested PCR. The amino-acid sequences used for degenerate primers were: *ScShh* first PCR, KQFIPNVA and AHIHCSV *ScShh* second, nested PCR, PNYNPDI and GFDWVYYE. The longer *ScShh* (348 bp) fragment was obtained from cDNA pools of stage 27 *S. canicula* embryos by RT-PCR using a degenerate forward primer and a *ScShh*-specific reverse primer (GRYEGKIT and TTCGTAGTAGACCCAGTC). The nucleotide sequences of *ScEn1*, *ScShh*, *ScTbx4*, *ScTbx5*, *ScTbx2*, *ScHand* and *Bmp4* cDNA are deposited in the GenBank database under the accession numbers: AF393834–AF393837, AY057890 and AY057891.

In situ hybridization

S. canicula embryos were removed from their egg casings and dissected from the yolk mass. Wholemount *in situ* hybridization on younger *S. canicula* embryos was carried out as previously described²³ and this is based on methods used for other vertebrate embryos²⁴. Older embryos were treated with dimethyl sulphoxide (DMSO) instead of proteinase K treatment by placing them in 2 ml of DMSO/methanol (1:1) on ice until they sank. Then 0.5 ml of 10% Triton X-100 (Sigma) in distilled water was added, and the embryos were incubated for an additional 20 minutes at room temperature^{25,26}. After washing in PBT (1% Tween 20 (Sigma) in PBS), embryos were hybridized with probes as described previously for chick embryos²⁴. Some whole-mount *in situ* samples were embedded in gelatin, and frozen sections were cut.

RT-PCR

RT-PCR was performed as previously described²⁷. The primers used for PCR amplification of *ScShh* (172 bp) were 5'-GAGCTGACAGGCTGATGACAC-3' and 5'-TGGTGATGTCCACAGCTCGGC-3'. The PCR cycle was at 96 °C for 20 seconds, 55 °C for 40 seconds and 72 °C for 1 minute for 32 cycles. Relative levels of transcripts were compared to levels of internal control using 18S ribosomal RNA primers (Ambion). Both *ScShh* and 18S rRNA primers were added into the same reaction solution.

Observation of cartilaginous pattern

S. canicula embryos to be stained for cartilage were fixed in 5% TCA (trichloroacetic acid), stained in 0.1% Alcian blue in 70% acid alcohol, dehydrated in ethanol and cleared in methyl salicylate.

Dil labelling

Dil (1,1-dioctadecyl-3,3,3',3'-tetramethylindo-carbocyanine perchloride; Molecular Probes; 3 mg ml⁻¹ in DMSO) was injected into chick wing-bud using a micropipette to label a small group of cells. Embryos were then incubated for 96 h and fixed in 4% paraformaldehyde in PBS. The average size of the initial Dil injected dot was 40–50 µm.

Received 2 July; accepted 17 December 2001.

- Thacher, J. K. Median and paired fins, a contribution to the history of vertebrate limbs. *Trans. Conn. Acad.* **3**, 281–310 (1877).
- Jarvik, E. in *Basic Structure and Evolution of Vertebrates* 109–131 (Academic, London, 1980).
- Moy-Thomas, J. A. The evolution of the pectoral fins of fishes and the tetrapod forelimb. *School Sci. Rev.* **36**, 592–599 (1936).
- Coates, M. I. The origin of vertebrate limbs. *Development* (1994 Suppl.), 169–180 (1994).
- Shu, D.-G. *et al.* Lower Cambrian vertebrates from South China. *Nature* **402**, 42–46 (1999).
- Neyt, C. *et al.* Evolutionary origins of vertebrate appendicular muscle. *Nature* **408**, 82–86 (2000).
- Balfour, F. M. The development of elasmobranch fishes. *J. Anat. Physiol. Lond.* **11**, 128–172 (1876).
- Ballard, W. W., Mellinger, J. & Lechenault, H. A series of normal stages for development of *Scyliorhinus canicula*, the lesser spotted dogfish (*Chondrichthyes: Scyliorhinidae*). *J. Exp. Zool.* **267**, 318–336 (1993).
- Altshuler, M., Clarke, J. D. & Tickle, C. Dorsal-ventral ectodermal compartments and origin of apical ectodermal ridge in developing chick limb. *Development* **124**, 4547–4556 (1997).
- Tanaka, M. *et al.* Apical ectodermal ridge induction by the transplantation of En-1-overexpressing ectoderm in chick limb bud. *Dev. Growth Differ.* **40**, 423–429 (1998).
- Tabin, C. & Laufer, E. *Hox* genes and serial homology. *Nature* **361**, 692–693 (1993).
- Coates, M. I. *Hox* genes, fin folds and symmetry. *Nature* **364**, 195–196 (1993).
- Tamura, K. *et al.* Evolutionary aspects of positioning and identification of vertebrate limbs. *J. Anat.* **199**, 195–204 (2001).
- Ruvinsky, I., Silver, L. M. & Gibson-Brown, J. J. Phylogenetic analysis of T-box genes demonstrates the importance of amphioxus for understanding evolution of the vertebrate genome. *Genetics* **156**, 1249–1257 (2000).
- Balinsky, B. I. Das Extremitätenfeld, seine Ausdehnung und Beschaffenheit. *Roux Arch. Dev. Biol.* **130**, 704–736 (1933).
- Cohn, M. J., Izpisua-Belmonte, J. C., Abud, H., Heath, J. K. & Tickle, C. Fibroblast growth factors induce additional limb development from the flank of chick embryo. *Cell* **80**, 739–746 (1995).
- Tanaka, M. *et al.* Distribution of polarizing activity and potential for limb formation in mouse and chick embryos and possible relationship to polydactyly. *Development* **127**, 4011–4021 (2000).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. *Sonic hedgehog* mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
- Ekker, S. C. *et al.* Patterning activities of vertebrate hedgehog proteins in the developing eye and brain. *Curr. Biol.* **5**, 44–55 (1995).
- Cohn, M. J. Giving limbs a hand. *Nature* **406**, 953–954 (2000).
- Tümpel, S. *et al.* Antero-posterior signaling in vertebrate limb development and stripes of *Tbx3* expression. *Dev. Biol.* (submitted).
- Kraus, P., Fraidenreich, D. & Loomis, C. A. Some distal limb structures develop in mice lacking Sonic hedgehog signaling. *Mech. Dev.* **100**, 45–58 (2001).

- Mazan, S., Jaillard, D., Baratte, B. & Janvier, P. *Otx1* gene-controlled morphogenesis of the horizontal semicircular canal and the origin of the gnathostome characteristics. *Evol. Dev.* **2**, 186–193 (2000).
- Wilkinson, D. G. *In Situ Hybridization: A Practical Approach* 75–83 (IRL Press/Oxford Univ. Press, Oxford, 1992).
- Kuratani, S., Ueki, T., Aizawa, S. & Hirano, S. Peripheral development of cranial nerves in a cyclostome, *Lampetra japonica*: morphological distribution of nerve branches and the vertebrate body plan. *J. Comp. Neurol.* **384**, 482–500 (1997).
- Schlosser, G. & Roth, G. Evolution of nerve development in frogs. I. The development of the peripheral nervous system in *Discoglossus pictus* (Discoglossidae). *Brain Behav. Evol.* **50**, 61–93 (1997).
- Münsterberg, A. E., Kitajewski, J., Bumcrot, D. A., McMahon, A. P. & Lassar, A. B. Combinatorial signaling by Sonic hedgehog and Wnt family members induces myogenic bHLH gene expression in the somite. *Genes Dev.* **9**, 2911–2922 (1995).
- Coates, M. I. Limb evolution. Fish fins or tetrapod limbs—a simple twist of fate? *Curr. Biol.* **5**, 844–848 (1995).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We are grateful to A. Wells for his assistance in maintenance of *S. canicula* embryos, S. Kuratani for information about *S. canicula* developmental studies before publication, S. Mazan for technical advice and *ScOtx1* and *ScOtx2* cDNA as positive control probes for establishing *in situ* hybridization methods and N. Helps for DNA sequencing. M.T. is supported by JSPS Postdoctoral Fellowships for Research Abroad, JSPS Research Fellowships for Young Scientists and the Inoue Research Award for Young Scientists. A.M. is supported by a Wellcome Trust research Career Development Award. C.T. is Foulerton Research Professor of The Royal Society.

Correspondence and requests for materials should be addressed to M.T. (e-mail: m.tanaka@dundee.ac.uk).

The cost of inbreeding in *Arabidopsis*

Carlos D. Bustamante[†], Rasmus Nielsen[‡], Stanley A. Sawyer[§], Kenneth M. Olsen^{||}, Michael D. Purugganan^{||} & Daniel L. Hartl^{*}

^{*} Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

[‡] Department of Biometrics, Cornell University, Ithaca, New York 14853-2801, USA

[§] Department of Mathematics, Washington University, St Louis, Missouri 63130, USA

^{||} Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695-7614, USA

Population geneticists have long sought to estimate the distribution of selection intensities among genes of diverse function across the genome. Only recently have DNA sequencing and analytical techniques converged to make this possible. Important advances have come from comparing genetic variation within species (polymorphism) with fixed differences between species (divergence)^{1,2}. These approaches have been used to examine individual genes for evidence of selection. Here we use the fact that the time since species divergence allows combination of data across genes. In a comparison of amino-acid replacements among species of the mustard weed *Arabidopsis* with those among species of the fruitfly *Drosophila*, we find evidence for predominantly beneficial gene substitutions in *Drosophila* but predominantly detrimental substitutions in *Arabidopsis*. We attribute this difference to the *Arabidopsis* mating system of partial self-fertilization, which corroborates a prediction of population genetics theory^{3–6} that species with a high frequency of inbreeding are less efficient in eliminating deleterious mutations owing to their reduced effective population size.

We analysed *Arabidopsis* data for 12 genes of diverse function for

[†] Present address: Department of Statistics, University of Oxford, Oxford OX1 3TG, UK.

which alleles were sequenced from two species. In each case, multiple alleles were sequenced from the partly self-fertilizing species *A. thaliana* and one allele from the closely related out-crossing species *A. lyrata*. For comparison, we analysed 34 genes from *D. melanogaster* and its sibling species *D. simulans*. Presentation of the data for each individual gene conventionally takes the form of a two-by-two table (Table 1), in which the number of nucleotide differences in each gene that do not change the encoded amino acid (synonymous) are categorized either as fixed within species but different between species (K_s) or as polymorphisms within one or both species (S_s). Similarly, the number of nucleotide differences in each gene that do change the encoded amino-acid (replacement) are categorized as either fixed (K_a) or polymorphic (S_a).

We use the term DPRS to refer to the layouts in Table 1 because, in clockwise order, the headings are divergence (D), polymorphism (P), replacement (R), and synonymous (S). Table 1 shows the means and s.d. (and range) of the counts in each cell for the genes from *Arabidopsis* and *Drosophila*. By itself, none of the individual DPRS tables is very illuminating as to whether amino-acid replacements (K_a) are selectively neutral relative to synonymous substitutions (K_s). In the absence of selection, the ratio of the expectations of K_s to S_s should equal the ratio of the expectations of K_a to S_a , so that selection is detected as a significant P value in a conventional test for homogeneity^{1,2}. Among the individual DPRS tables, when the P values are corrected for multiple comparisons, there is only one gene in the *Arabidopsis* data that is significant at the 5% level, and only two genes in the *Drosophila* data that are significant at the 5% level (Fisher exact tests, data not shown).

Although few of the individual DPRS tables are significant, each contains information about the selective forces impinging on the amino-acid replacements. A theoretical framework for extracting this information derives from considerations of the equilibrium flux of fixations and limiting probability densities of nucleotide substitutions affected by mutation, selection and random genetic drift taking place simultaneously and independently at multiple sites in a DNA sequence². In this framework, known as a Poisson random field (PRF), the magnitude of each cell observed in a DPRS table is an independent Poisson random variable whose expected value² is given by the corresponding equation in Table 1 (bottom). The symbols n and m are the number of alleles sequenced from each of the species being compared. The quantity $\frac{1}{2}\theta_s$ is the expected number of synonymous mutations that occur in the gene in the entire population in any generation; $\frac{1}{2}\theta_a$ is the corresponding quantity for non-synonymous mutations (amino-acid replacements), excluding those mutations that are so deleterious that they have a negligible chance of becoming polymorphic or fixed. The parameter of greatest present interest is γ , which is the selection intensity in favour of (if $\gamma > 0$) or against (if $\gamma < 0$) amino acid replacements; γ is scaled according to the haploid effective population number N_e ; hence, γ equals the selection coefficient multi-

plied by N_e . (The haploid effective size is twice the diploid effective size.) Although synonymous sites in some genes are known to be under selection for optimal codon usage, this effect is usually small^{7,8} and is not explicitly taken into account in the present formulation. Consequently, γ can be thought of as the magnitude of selection affecting amino-acid replacements relative to that affecting synonymous substitutions. Because our main interest is in variation in the intensity of selection among genes, we assume that all amino-acid polymorphisms and fixed differences in the same gene have a common γ , but that γ can differ from one gene to the next. The fourth parameter, t , is the number of generations since the species diverged from a common ancestor, again expressed as a multiple of N_e . The three functions in Table 1 (bottom) are defined as

$$L(n) = \sum_{i=1}^{n-1} \frac{1}{i}$$

$$F(n) = \int_0^1 \frac{1 - x^n - (1 - x)^n}{1 - x} \frac{1 - e^{-2\gamma x}}{2\gamma x} dx$$

$$G(n) = \int_0^1 (1 - x)^{n-1} \frac{1 - e^{-2\gamma x}}{2\gamma x} dx$$

Although each DPRS table has four parameters (θ_s , θ_a , γ and t) and four observations (K_a , K_s , S_a and S_s), the divergence time t is a common parameter. This implies that each gene contributes valuable information about the distribution of γ among genes. We therefore chose an analytical method that borrows information from all the genes to make inferences about the magnitude of selection for any individual gene. This approach greatly increases the power and accuracy of the inferences regarding selection.

A suitable framework is provided by the hierarchical Bayesian model described in Box 1 (refs 9, 10). For each species pair, we assume that the magnitude of γ for each gene is drawn randomly and independently from a normal distribution with mean μ and standard deviation σ . The hierarchical structure is achieved by assuming that μ and σ are themselves random variables. On the basis of this model we estimate the probability distribution of μ given the observed data and show that this distribution for genes in *Arabidopsis* is significantly different from that for genes in *Drosophila*.

The equation for the posterior distribution $\pi(\gamma, t, \theta, \mu, \sigma | D)$ given in the box is analytically intractable. Nevertheless, parameter estimates based on the posterior distribution are possible. The approach is first to define a Markov chain that has $\pi(\gamma, t, \theta, \mu, \sigma | D)$ as its stationary distribution. The means, modes, variances and other quantities for individual parameters are approximated by sampling from one or more long trajectories of this Markov chain with different starting positions. The Markov chain is defined by a sampling method, called Markov Chain Monte Carlo (MCMC)¹¹, whose properties result in convergence to $\pi(\gamma, t, \theta, \mu, \sigma | D)$. By

Table 1 Observed and expected numbers of amino-acid changes that are fixed between species or polymorphic within species

Amino-acid change	Divergence	Polymorphism
<i>Arabidopsis</i>		
Synonymous	$K_s = 26.7 \pm 14.0$ (range 1–55)	$S_s = 9.5 \pm 7.1$ (range 1–25)
Replacement	$K_a = 11.2 \pm 6.6$ (range 3–26)	$S_a = 9.3 \pm 6.5$ (range 2–19)
<i>Drosophila</i>		
Synonymous	$K_s = 18.5 \pm 12.5$ (range 1–49)	$S_s = 17.4 \pm 17.2$ (range 0–69)
Replacement	$K_a = 10.9 \pm 16.3$ (range 0–75)	$S_a = 5.7 \pm 8.9$ (range 0–37)
Poisson random-field expected values		
Synonymous	$E(K_s) = \theta_s \left(t + \frac{1}{m} + \frac{1}{n} \right)$	$E(S_s) = \theta_s [L(m) + L(n)]$
Replacement	$E(K_a) = \theta_a \left(\frac{2\gamma}{1 - e^{-2\gamma}} \right) [t + G(m) + G(n)]$	$E(S_a) = \theta_a \left(\frac{2\gamma}{1 - e^{-2\gamma}} \right) [F(m) + F(n)]$

Arabidopsis data are means $\pm$ s.d. among 12 genes from *A. thaliana* and *A. lyrata*; the number of alleles ranges from 14 to 21 (mean 17.6), and the length of coding region sequenced from 531 to 1,674 base pairs (bp) (mean 935 bp). *Drosophila* data are means $\pm$ s.d. among 34 genes from *D. melanogaster* and *D. simulans*; the number of alleles ranges from 5 to 62 (mean 10.8) and the length of coding region sequenced from 270 to 7,683 bp (mean 1172 bp).

custom, a relatively large number of iterations at the beginning of each realization of the Markov chain is disregarded as a 'burn-in' period. This is a protection against possible bias caused by the starting conditions.

For each trajectory in the MCMC simulations, we chose arbitrary initial values for μ and σ , the selection parameters, and the divergence time. Then, for each gene in turn, given its value of γ , a value for θ_a was drawn from a gamma distribution

$$\Gamma\left\{\alpha + K_a + S_a, \beta + \left(\frac{2\gamma}{1 - e^{-2\gamma}}\right)[t + G(m) + G(n) + F(m) + F(n)]\right\}$$

which is the posterior distribution of θ_a given the data and all other relevant parameters. Similarly, a value for θ_s was chosen from its posterior distribution

$$\Gamma[\alpha + K_s + S_s, \beta + t + 1/m + 1/n + L(m) + L(n)]$$

In both cases we choose α and β to be close to 0 and hence uninformative.

Iterative updating of the selection parameters proceeds by Metropolis sampling¹² as follows. For each gene in turn, choose a proposed new value of γ —call it γ' —from a narrow uniform distribution centred on γ . Evaluate the Poisson means in Table 1 (bottom) and calculate the likelihood of the data for the value γ' (using the current value of θ_a) and multiply by the prior probability of γ' calculated from the normal distribution with mean μ and variance σ^2 . This is essentially the posterior probability for γ' , and if it is greater than the posterior probability for γ , set $\gamma = \gamma'$; otherwise set $\gamma = \gamma'$ only if a uniform random number in $[0, 1]$ is less than the ratio of the posterior probabilities. The mutation parameters for each gene in turn are updated by Gibbs sampling¹³, which is implemented by choosing new values from the updated gamma distributions. Once all the selection and mutation parameters have been updated, the divergence time is updated by Metropolis sampling in a manner analogous to the updating of the γ values. To update μ and σ , first calculate the sample mean ($\bar{\gamma}$) and variance (s_γ^2) of the updated γ values. Then sample σ^2 from its posterior distribution, which is inverse-gamma-distributed with parameters that depend on s_γ^2 and the number of genes k . Finally, update μ , which is normally distributed with mean and variance

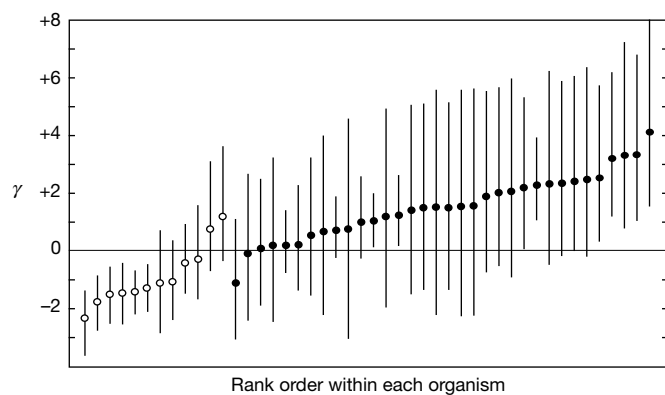


Figure 1 Means of the posterior distributions of the selection parameters γ (dots) and the 95% credible intervals (vertical lines), in rank order according to the mean, for the genes in *Arabidopsis* (open circles) and *Drosophila* (filled circles). From left to right the *Arabidopsis* genes are AP3, PgiC, Pl, AP1, ChiA, CAL, TFL1, FAH1, Adh1, F3H, LFY and CHI, and the *Drosophila* genes are per, Pgi, Acp26Ab, Adh, Est-6, pn, boss, CecA2, Anon1G5, thy, ref(2)p, Anon1E9, Tpi, Anon1A3, CecA1, slgA, Pp4-19C, AnnX, Cdic, Bap, shakB, gld, Dhc-Yh3, runt, Acp26Aa, z, 6-Pgd, Yp2, Rh3, ase, Acp29AB, cta, G6pd and ci. Sources of the data are referenced in the Supplementary Information.

that depend on $\bar{\gamma}$, σ^2 and k . This procedure of updating the γ , θ , t , μ and σ^2 is then repeated to form the Markov chain of parameter values whose stationary distribution is given by $\pi(\gamma, t, \theta, \mu, \sigma|D)$. The range of the uniform distributions used to update the selection parameters and the divergence time were chosen by preliminary trials to yield an acceptance rate of about 50%.

For each species pair, we ran five independent chains from overdispersed starting points for 10^6 iterations each. The first 10^5 iterations were treated as burn-in and disregarded. The chains converged and mixed very fast, yielding estimated scale reduction factors¹⁴ close to 1 (1.00001 for *Arabidopsis* and 0.99994 for

Box 1

Hierarchical bayesian analysis of polymorphism and divergence

In bayesian statistics, assumptions about the forms of the underlying distributions (prior distributions) of parameters are combined with current data by the use of a likelihood function. The result is the posterior distribution of the parameters, given the data. In our case, the parameters of interest include a vector γ of continuous-time selection coefficients (Malthusian fitnesses). We assume that the individual γ values, one for each gene, are independent samples from a normal distribution of selection coefficients with some unknown mean μ and unknown variance σ^2 . The values of μ and σ^2 are of interest, as are the species divergence time (t) and the vector of mutation parameters (θ). For k DPRS tables, there are k selection coefficients and $2k$ mutation parameters. We adopt the bayesian approach because it enables estimates with specified degrees of confidence even for correlated parameters, it allows the sharing of data across DPRS tables to estimate the divergence time, and it provides a powerful computational tool. The goal is to use the observed data (D) to make inferences about the posterior distribution of the parameters ($\gamma, t, \theta, \mu, \sigma$). This posterior distribution is symbolized as $\pi(\gamma, t, \theta, \mu, \sigma|D)$. The bayesian formulation of our model is given formally by

$$\pi(\gamma, t, \theta, \mu, \sigma|D) = \frac{P(D|\gamma, t, \theta)f(\gamma|\mu, \sigma)g(\mu|\sigma)h(\sigma)p(\theta)q(t)}{\int \int \int \int \int [P(D|\gamma, t, \theta)f(\gamma|\mu, \sigma)g(\mu|\sigma)h(\sigma)p(\theta)q(t)]d\gamma dt d\theta d\mu d\sigma}$$

Properties of individual parameters in the posterior distribution are used for estimates and credible intervals (bayesian confidence intervals). The value of the expression $P(D|\gamma, t, \theta)$ is calculated from the independent Poisson distributions whose means are given in Table 1 (bottom), and the other functions are the assumed prior probability distributions. For example, $f(\gamma|\mu, \sigma)$ is a normal distribution with mean μ and variance σ^2 , whereas $g(\mu|\sigma)$ and $h(\sigma)$ embody the hierarchical feature that μ and σ are themselves treated as random variables, with $g(\mu|\sigma)$ a normal distribution and $h(\sigma)$ such that $1/\sigma^2$ is a gamma distribution. These choices are motivated by the facts that the sample mean of independent observations from a normal distribution, given σ , is itself normal, and that the distribution of the sample variance is gamma. We also assume that the prior $p(\theta)$ for mutation is gamma, because with this choice the distribution of these parameters in the posterior distribution has the same form as in the prior distribution, and that the prior $q(t)$ for the divergence time is uniform. These priors are all chosen to be 'uninformative' in the sense that the parameters in the posterior distribution are determined primarily by the current data and not by the characteristics of the prior distribution.

Many simplifications result from this bayesian formulation as applied to DPRS tables. For example, because the DPRS tables are independent, the multivariate distribution $f(D|\gamma, t, \theta)$ can be written as a product of univariate distributions $f(D_i|\gamma_i, t, \theta)$, where the subscript i indexes the gene ($i = 1, 2, \dots, k$). Likewise, the multivariate normal distribution $g(\gamma|\mu, \sigma)$ can be written as a product of univariate normal distributions of the form $g(\gamma_i|\mu, \sigma)$. Because mutation enters the equations in the DPRS tables multiplicatively, updating the mutation parameters conditional on γ and t is straightforward. These simplifications also imply that many of the parameters become uncorrelated in their conditional distributions.

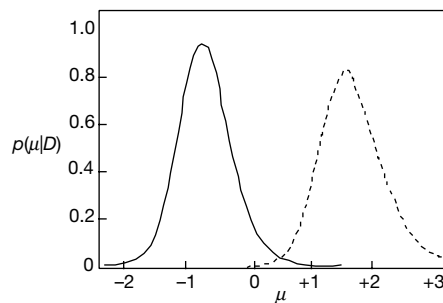


Figure 2 Estimated distribution of the mean value of the selection parameter (γ) across genes in *Arabidopsis* (solid line) and *Drosophila* (dashed line).

Drosophila). After the burn-in, each chain was sampled every 100 generations, yielding 45,000 sample points for each species pair.

Using the results of MCMC, we estimated the mean of the posterior distribution of the selection intensity γ for each gene. These are shown in rank order in Fig. 1, along with the 95% credible intervals. For the 12 *Arabidopsis* genes, 10 have their mean $\gamma < 0$, whereas for the 34 *Drosophila* genes, 32 have their mean $\gamma > 0$. Among the 12 *Arabidopsis* genes, 6 credible intervals are entirely negative (do not overlap 0), and among the 34 *Drosophila* genes, 9 are entirely positive. The detection of non-zero selection parameters for so many genes contrasts with conventional significance tests and demonstrates the power of the bayesian analysis.

The posterior distribution of μ also shows a significant difference between the organisms (Fig. 2). For *Arabidopsis* genes, most of the posterior distribution for the average selective effect of polymorphic or fixed amino-acid replacements has $\mu < 0$, and in fact the probability that $\mu > 0$ is about 0.03. In contrast, for *Drosophila* genes, most of the posterior distribution has $\mu > 0$, and indeed the probability that $\mu < 0$ is about 0.0004. The overall probability that μ for *Arabidopsis* is actually greater than μ for *Drosophila* is $P = 0.001$, which we estimated by comparing the 45,000 sample values for each data set in all possible pairs.

We conclude from Figs 1 and 2 that the average amino-acid replacement that is polymorphic or fixed in *Drosophila* is beneficial, whereas the average amino-acid replacement that is polymorphic or fixed in *Arabidopsis* is slightly deleterious. This result is consistent with a recent decrease in effective population size in *A. thaliana*, which is predicted^{3–6} to result from its mode of reproduction largely by self-fertilization¹⁵. *A. lyrata* reproduces by outcrossing¹⁶, and phylogenetic evidence indicates that partial selfing is a derived trait¹⁷.

If *A. thaliana* and *A. lyrata* were to share many ancestral polymorphisms, some of these would be scored as fixed differences. However, these species show substantial sequence divergence (Table 1, top) and are estimated to have diverged 5–6 Myr ago¹⁷. Our analysis indicates that the 95% credible interval for the divergence time t between the species, in multiples of the haploid effective population size, is 6.9–11.2. Considering that a newly arisen neutral mutation that is destined to become fixed has an average fixation time of $t = 2$ (as a multiple of the haploid effective size), extensive shared polymorphism seems unlikely. For the *Drosophila* species, which diverged 1–3 Myr ago¹⁸, the 95% credible interval for t is 3.7–4.8.

Tight linkage of the nucleotide sites within a gene is also an issue. However, computer simulations indicate that the effect of linkage is to bias the estimates of γ towards 0 (data not shown). Hence, linkage would tend to make the differences in the selection parameters between *Arabidopsis* and *Drosophila* seem somewhat smaller than they really are.

Beyond the present application, hierarchical bayesian analysis affords a theoretical framework for a science of evolutionary genomics. In principle it could discriminate between the evolutionary forces that affect proteins of diverse function, identifying effects specific to sex, tissue and developmental stage of expression, cellular

location, three-dimensional structure and mechanism of action. The analysis could also sort out the evolutionary forces affecting upstream, downstream, intronic and intergenic regions of genes. Even in this era of high-throughput genomic sequencing, the acquisition of genome-wide polymorphism-divergence data presents a formidable challenge.

Received 6 August 2001; accepted 3 January 2002.

- McDonald, J. H. & Kreitman, M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652–654 (1991).
- Sawyer, S. A. & Hartl, D. L. Population genetics of polymorphism and divergence. *Genetics* **132**, 1161–1176 (1992).
- Charlesworth, D., Morgan, M. T. & Charlesworth, B. The effect of deleterious mutations on neutral molecular variation. *J. Hered.* **84**, 321–325 (1993).
- Kondrashov, A. S. Muller's ratchet under epistatic selection. *Genetics* **136**, 1469–1473 (1994).
- Caballero, A. & Santiago, E. Response to selection from new mutation and effective size of partially inbred populations. I. Theoretical results. *Genet. Res.* **66**, 213–225 (1995).
- Wang, J. L., Hill, W. G., Charlesworth, D. & Charlesworth, B. Dynamics of inbreeding depression due to deleterious mutations in small populations: Mutation parameters and inbreeding rate. *Genet. Res.* **74**, 165–178 (1999).
- Hartl, D. L., Moriyama, E. N. & Sawyer, S. A. Selection intensity for codon bias. *Genetics* **138**, 227–234 (1994).
- Akashi, H. Inferring weak selection from patterns of polymorphism and divergence at 'silent' sites in *Drosophila* DNA. *Genetics* **139**, 1067–1076 (1995).
- Gelman, A., Carlin, J. S., Stern, H. S. & Rubin, D. B. *Bayesian Data Analysis* (Chapman & Hall, London, 1997).
- Carlin, B. P. & Louis, T. A. *Bayes and Empirical Bayes Methods for Data Analysis* (Chapman & Hall, London, 2000).
- Gilks, R., Richardson, S. & Spiegelhalter, D. J. *Markov Chain Monte Carlo in Practice* (Chapman & Hall, London, 1996).
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1091 (1953).
- Geman, S. & Geman, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.* **6**, 721–741 (1984).
- Gelman, A. & Rubin, D. B. Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7**, 457–511 (1992).
- Abbott, R. J. & Gomes, M. F. Population genetic structure and the outcrossing rate of *Arabidopsis thaliana*. *Heredity* **62**, 411–418 (1989).
- Savolainen, O., Langley, C. H., Lazzaro, B. P. & Freville, H. Contrasting patterns of nucleotide polymorphism at the *Adh* locus in the outcrossing *Arabidopsis lyrata* and the selfing *Arabidopsis thaliana*. *Mol. Biol. Evol.* **17**, 645–655 (2000).
- Kusaba, M. et al. Self-incompatibility in the genus *Arabidopsis*: Characterization of the *S* locus in the outcrossing *A. lyrata* and its autogamous relative *A. thaliana*. *Plant Cell* **13**, 627–643 (2001).
- Caccone, A., Amato, G. D. & Powell, J. R. Rates and patterns of scdDNA and mtDNA divergence within the *Drosophila melanogaster* subgroup. *Genetics* **118**, 671–683 (1988).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>) and on the authors' website (<http://www.oeb.harvard.edu/hartl/lab>).

Acknowledgements

We thank D. Weinreich and D. Rand for providing the *Drosophila* data, and A. Kondrashov for numerous suggestions for improving the presentation. This work was supported by grants from the US Public Health Service, the US National Science Foundation, Howard Hughes and Marshall Sherfield Fellowships to C.D.B., and an Alfred P. Sloan Foundation Young Investigator Award to M.D.P.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.L.H. (e-mail: dhartl@oeb.harvard.edu).

Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation *in vivo*

Dominic M. Walsh*, Igor Klyubin†, Julia V. Fadeeva*, William K. Cullen†, Roger Anwyl‡, Michael S. Wolfe*, Michael J. Rowan† & Dennis J. Selkoe*

* Department of Neurology, Harvard Medical School and Center for Neurologic Diseases, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

† Departments of Pharmacology and Therapeutics and ‡ Physiology, Trinity College, Dublin 2, Ireland

Although extensive data support a central pathogenic role for amyloid β protein (A β) in Alzheimer's disease¹, the amyloid hypothesis remains controversial, in part because a specific neurotoxic species of A β and the nature of its effects on synaptic function have not been defined *in vivo*. Here we report that natural oligomers of human A β are formed soon after generation of the peptide within specific intracellular vesicles and are subsequently secreted from the cell. Cerebral microinjection of cell medium containing these oligomers and abundant A β monomers but no amyloid fibrils markedly inhibited hippocampal long-term potentiation (LTP) in rats *in vivo*. Immunodepletion from the medium of all A β species completely abrogated this effect. Pretreatment of the medium with insulin-degrading enzyme, which degrades A β monomers but not oligomers, did not prevent the inhibition of LTP. Therefore, A β oligomers, in the absence of monomers and amyloid fibrils, disrupted synaptic plasticity *in vivo* at concentrations found in human brain and cerebrospinal fluid. Finally, treatment of cells with γ -secretase inhibitors prevented oligomer formation at doses that allowed appreciable monomer production, and such medium no longer disrupted LTP, indicating that synaptotoxic A β oligomers can be targeted therapeutically.

Fibrillar (but not monomeric) forms of A β akin to those present in the amyloid plaques of Alzheimer's disease are neurotoxic in

culture^{2,3}. However, relatively weak correlations between fibrillar plaque density and severity of dementia are found in Alzheimer's diseased brains^{4,5}, whereas correlations between soluble A β levels and the extent of synaptic loss and cognitive impairment are stronger^{6,7}. SDS-stable A β oligomers (of relative molecular masses $M_r \sim 8,000$ and $\sim 12,000$) have been detected by western blotting in the buffer-soluble fraction of Alzheimer's diseased cortex⁷. These are strikingly similar to soluble, SDS-stable A β oligomers produced by certain cultured cells^{8–10}. We previously confirmed the latter species as A β oligomers by amino-terminal radiosequencing and precipitation with carboxy-terminal-specific antibodies^{8,10}. The levels of these oligomers are specifically augmented by expressing Alzheimer's disease-causing mutations in amyloid precursor protein (APP) or presenilin that increase A β 42 production¹¹, supporting their pathological relevance. Importantly, young APP transgenic mice undergo synaptic, electrophysiological and behavioural changes before any amyloid plaque formation^{12,13}, but the nature of the responsible A β species in the brain cannot be specifically defined.

To assess the subcellular origin of naturally occurring A β oligomers secreted by cells expressing human APP, we prepared microsomes from 7PA2 Chinese hamster ovary (CHO) cells, which stably express the V717F Alzheimer's disease mutation in APP₇₅₁ (an APP:isoform 751 amino acids in length)^{8,10}. A highly sensitive immunoprecipitation/western blot method¹⁰ revealed both monomers and oligomers of A β in the isolated, washed microsomes, and the A β dimers in the microsomes co-migrated with those in the conditioned medium (CM) (Fig. 1a). Previous trypsinization of the intact cells did not alter the amounts of the various A β assemblies recovered in the microsomes (Fig. 1a, lanes 7 versus 8), supporting their intracellular origin. The co-migrating dimers and trimers present in CM reacted with antibodies specific to the C termini of A β 40 and A β 42 (Fig. 1a, lanes 3 and 4). The oligomeric bands were not altered by pre-treatment with 8 M urea or 50% formic acid (Fig. 1b). Previously, radiosequencing of these monomer, dimer and trimer bands in 7PA2 CM confirmed that they are composed of A β ⁸.

Extensive immunoelectron microscopy, which readily detects synthetic A β protofibrils in the CM of cultured cells¹⁴, revealed no such filamentous assemblies in 7PA2 CM (not shown). Ultracentrifugation of the CM at a force ($100,000g \times 3$ h) that sediments synthetic A β fibrils and protofibrils did not bring down any A β

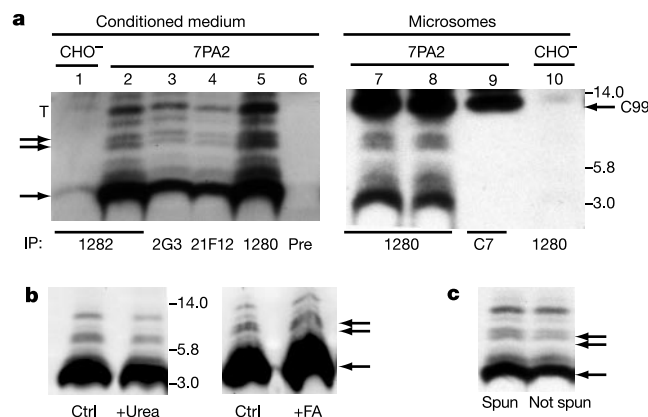


Figure 1 SDS-stable oligomers of A β are produced intracellularly and secreted. **a**, 7PA2 (lanes 2–6) and CHO⁻ (lane 1) conditioned medium (CM) were immunoprecipitated (IP) with antibodies or preimmune serum (Pre) and western blotted with monoclonal antibody 6E10. Microsomes from either 7PA2 (lanes 7–9) or CHO⁻ (lane 10) cells were incubated with (lane 7) or without (lanes 8, 10) trypsin, lysed and precleared with APP C-terminal antiserum, C7 (pellet shown, lane 9). Precleared, lysates were then precipitated with R1280 (lanes 7, 8 and 10). The A β trimer in microsomal

lysates is obscured by residual APP C-terminal fragment, C99 (compare lanes 7–9 with lanes 2–5). Arrow, monomeric A β ; double arrow, dimeric A β ; T, trimer. **b**, Incubation of the R1282 precipitate of 7PA2 CM in either 8 M urea or 50% formic acid did not alter oligomers. **c**, Centrifugation of CM at $100,000g \times 3$ h did not alter oligomer amounts. Relative molecular masses ($M_r \times 1,000$) are shown on the right-hand side of gel lanes in **a**, **b**, Figs 2, 3c–e and 4d, h.

assemblies, and the amount of oligomers in the supernatant was unaltered (Fig. 1c).

Next, we fractionated total 7PA2 microsomes on discontinuous iodixanol gradients¹⁵ and compared the distribution of the oligomers with those of organelle marker proteins. A β immunoreactive bands at M_r 4K, 6K and 8K–9K were principally present in fractions 4–7, which were enriched in Golgi-type vesicles (syntaxin-6-positive) and recycling endosomes (transferrin-receptor-positive) but contained only low levels of endoplasmic reticulum proteins (Grp78 and calnexin) (Fig. 2 and not shown). The microsomal A β species co-migrated with the previously confirmed oligomers in the CM (Fig. 2d) and were unreactive with an APP C-terminal antibody (C7) (Fig. 2d) or preimmune serum (not shown). Immunoprecipitation of the subcellular vesicles with antibodies to the free C termini of A β 40 (2G3) and A β 42 (21F12) revealed a faint A β species of M_r 12K, a putative trimer, in addition to monomers and dimers (not shown). These were again principally localized to iodixanol gradient fractions 4–7, in which A β has previously been shown to be generated¹⁵. These data suggest that A β oligomerization is initiated soon after the generation of A β in discrete intracellular vesicles.

To determine whether these cell-derived oligomers of human A β have biological activity *in vivo*, we examined the effect of 7PA2 CM on hippocampal LTP, a measure of synaptic plasticity that is exquisitely sensitive to disruption by synthetic A β ¹⁶. Long-term potentiation of excitatory synaptic transmission in the CA1 area of anaesthetized rats was induced by high-frequency stimulation (HFS). Microinjection of 7PA2 CM (1.5 μ l intracerebroventricular, i.c.v.) completely blocked LTP at 3 h post-stimulation; only a waning potentiation lasting <90 min was elicited (Fig. 3a, g). In contrast, a robust LTP that was stable for over 3 h was induced after identical infusion of CM (1.5 μ l) from untransfected CHO[−] sister cultures (Fig. 3a, g). The LTP block was not related to any decrease in baseline transmission, because 7PA2 CM did not alter synaptic responses (EPSPs) in the absence of HFS (Fig. 3a, g).

The CM of CHO[−] and 7PA2 cells should be indistinguishable except for the presence in the latter of the secretory products of human APP metabolism: APP_s, A β and p3 (A β 17–40/42). To determine if A β caused the LTP block, we compared the effects of

7PA2 CM before and after immunodepletion with R1282, a high-titre polyclonal A β antibody. R1282 immunodepletion completely prevented the block of LTP (Fig. 3b, g). Western blotting confirmed that R1282 efficiently precipitated both A β monomers and oligomers; a second immunoprecipitation brought down very little additional protein (Fig. 3c). By enzyme-linked immunosorbent assay (ELISA), the R1282 precipitation decreased total A β in CM from $5,475 \pm 897$ to $1,253 \pm 24$ pg ml^{−1} (mean $\pm$ s.d., $n = 3$). APP_s was poorly precipitated by R1282 (not shown). Correspondingly, the amounts of APP_s remaining in CM before and after R1282 immunoprecipitation were indistinguishable (Fig. 3d). The abrogation of the LTP block by R1282 was not due to a nonspecific effect of the immunodepletion or to an undetected decrease in APP_s concentration, because immunodepletion of 7PA2 CM with a polyclonal antibody (B5) to APP_s had no effect (Fig. 3b, g). The addition of R1282 to CHO[−] CM likewise did not affect LTP induction (Fig. 3g).

Because neither the A β ELISA nor the immunoprecipitation/western blot assay recognizes p3 (ref. 10), we addressed the possibility that p3 in the 7PA2 CM mediated the observed block of LTP. Synthetic human p3 added to CHO[−] CM at a concentration (11.5 nM) well in excess of that found in 7PA2CM⁸ did not affect LTP (Fig. 3h).

Next, we asked whether the block of LTP was due to the abundant A β monomers or to the less abundant but invariably present SDS-stable A β oligomers. To do so, we took advantage of the specificity of insulin-degrading enzyme (IDE) to selectively degrade and remove A β monomers from CM (Fig. 3e)¹⁷. 7PA2 CM was divided into three equal aliquots: one was held at 4 °C, another was incubated at 37 °C for 12 h, and the third was incubated with purified IDE at 37 °C for 12 h. The latter incubation caused complete loss of detectable A β monomer, while leaving the A β oligomers virtually unchanged (Fig. 3e, lane 2). The CM aliquot incubated without IDE underwent only a slight decrease in A β monomer (Fig. 3e, lane 3). Both the second and third aliquots still caused a block of LTP at 3 h (Fig. 3f, h) that was indistinguishable from that of the first (untreated) aliquot. That IDE itself blocked LTP was excluded when (1) injection of plain medium spiked with IDE did not alter LTP (Fig. 3h), and (2) infusion of IDE-treated 7PA2 CM from which the His-tagged IDE had been removed by metal affinity chromatography still blocked LTP (Fig. 3h, IDE plus MAC). We conclude that naturally secreted, soluble human A β oligomers are specifically responsible for the block of hippocampal LTP by 7PA2 CM.

Because LTP was selectively blocked by A β oligomers in the absence of monomers, protofibrils or fibrils, we pharmacologically prevented oligomer formation in living cells while retaining substantial monomer production and assessed whether this would preserve normal LTP. We synthesized two structurally dissimilar γ -secretase inhibitors that dose-dependently inhibited A β secretion by 7PA2 cells: DAPM (half-maximal inhibitory concentration, IC₅₀ ~10 nM) and MWIII-20 (IC₅₀ ~10 μ M) (Fig. 4a, b). After treatment of cells for 6 h with 10 μ M MWIII-20, the dimer (M_r 8K) was only faintly detectable by immunoprecipitation/western blot, whereas ~60% of the monomer signal remained (Fig. 4c, e). For MWIII-20, the IC₅₀ for monomers by this method was ~30 μ M, whereas that for dimers was ~10-fold lower. With the more potent DAPM, the IC₅₀ for monomers was ~0.1 μ M, whereas that for dimers was ~20-fold lower. Thus, at 0.1 μ M, DAPM decreased monomers, but there was a relatively greater loss of dimers, which were now barely or not detectable (Fig. 4d). The percentage decrease of dimer at each concentration was greater than that of monomer in all experiments, and these two values were significantly different at the higher concentrations ($P < 0.005$) (Fig. 4e, f, asterisks). On the basis of these results, we treated 7PA2 cells for 6 h with 0.1 μ M DAPM, confirmed the marked decrease in oligomer levels in this CM (Fig. 4h) and then microinjected an aliquot (1.5 μ l i.c.v.) into rats. High-frequency stimulation produced a robust LTP which was

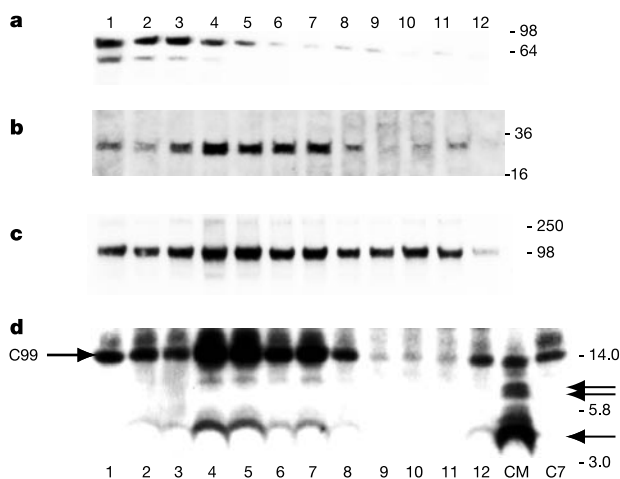


Figure 2 SDS-stable oligomers are formed soon after generation of A β . Microsomes fractionated on discontinuous iodixanol gradients¹⁵ were lysed and blotted for Grp78 (**a**), syntaxin 6 (**b**) and transferrin receptor (**c**). The remainder of each fraction was precleared with C7-protein A (as in Fig. 1) and then precipitated with R1280, and A β species blotted with 6E10 (**d**). CM, medium from the same cells. C7, aliquot of fraction 4, precipitated with C7. Arrow, monomeric A β ; double arrow, dimeric A β . Data are typical of four fractionations.

fully maintained for over 180 min (Fig. 4g, red symbols). This LTP was indistinguishable from that observed after injection of CHO⁻ CM spiked with the same or greater concentrations of DAPM (1 μ M, $144 \pm 8\%$, $n = 3$). To control further for the presence of the inhibitor and the fact that treatment with DAPM reduced A β monomer production by $\sim 40\%$, 7PA2 CM was diluted 1:1 with plain DMEM, and DAPM spiked in (to 0.1 μ M) just before injection. In contrast to the results obtained using CM from DAPM-treated cells, injection of this diluted 7PA2 CM spiked with DAPM caused a block of LTP typical of that observed previously ($P < 0.05$, Fig. 4g, blue symbols).

These experiments allow us to attribute an inhibition of hippocampal LTP *in vivo* specifically to oligomers, not monomers or fibrils, of naturally secreted human A β . The ability to collect heterogeneous human A β species produced naturally by living cells, combined with the property of IDE to degrade A β monomers but not oligomers, enabled us to specify that SDS-stable oligomers of A β consistently block hippocampal LTP. In addition to the advantages of this approach over treatments using synthetic A β , it provides specificity about the form of A β responsible for biological activity that is not achievable in APP transgenic mouse brains. The work demonstrates for the first time that a biochemically defined oligomeric assembly of naturally secreted human A β alters hippo-

campal synaptic efficacy at physiological levels. Our findings strongly support the emerging hypothesis that soluble A β oligomers are the principal effectors of the synaptic dysfunction and loss that characterize Alzheimer's disease^{7,12-14,18-20}. Importantly, the CM we microinjected contained total human A β concentrations (mean $\pm$ s.d.: $3,223 \pm 1,217$ pg ml⁻¹; $n = 9$) very similar to those we previously measured in normal human cerebrospinal fluid ($4,003 \pm 1,873$ pg ml⁻¹, $n = 32$)¹⁰.

We provide evidence that after their genesis in intracellular vesicles, A β monomers form dimers, trimers and perhaps higher oligomers (Figs 1 and 2), that at least a portion of these oligomers is highly stable, and that some are subsequently secreted (Fig. 1b). The secreted oligomers can interact with neurons *in vivo*, altering their normal physiology (Fig. 3). Analogous effects may underlie the subtle synaptic changes and impairment of learning and memory documented in young APP transgenic mice^{12,13,20-21} and in early Alzheimer's disease patients themselves. Our findings do not rule out an additional role for subsequent morphological lesions that characterize Alzheimer's disease and certain transgenic mouse models²².

Synthetic A β oligomers block rat hippocampal LTP¹⁸ but involve substantially higher doses of single synthetic A β species. In transgenic mice overexpressing human APP, some studies report failure

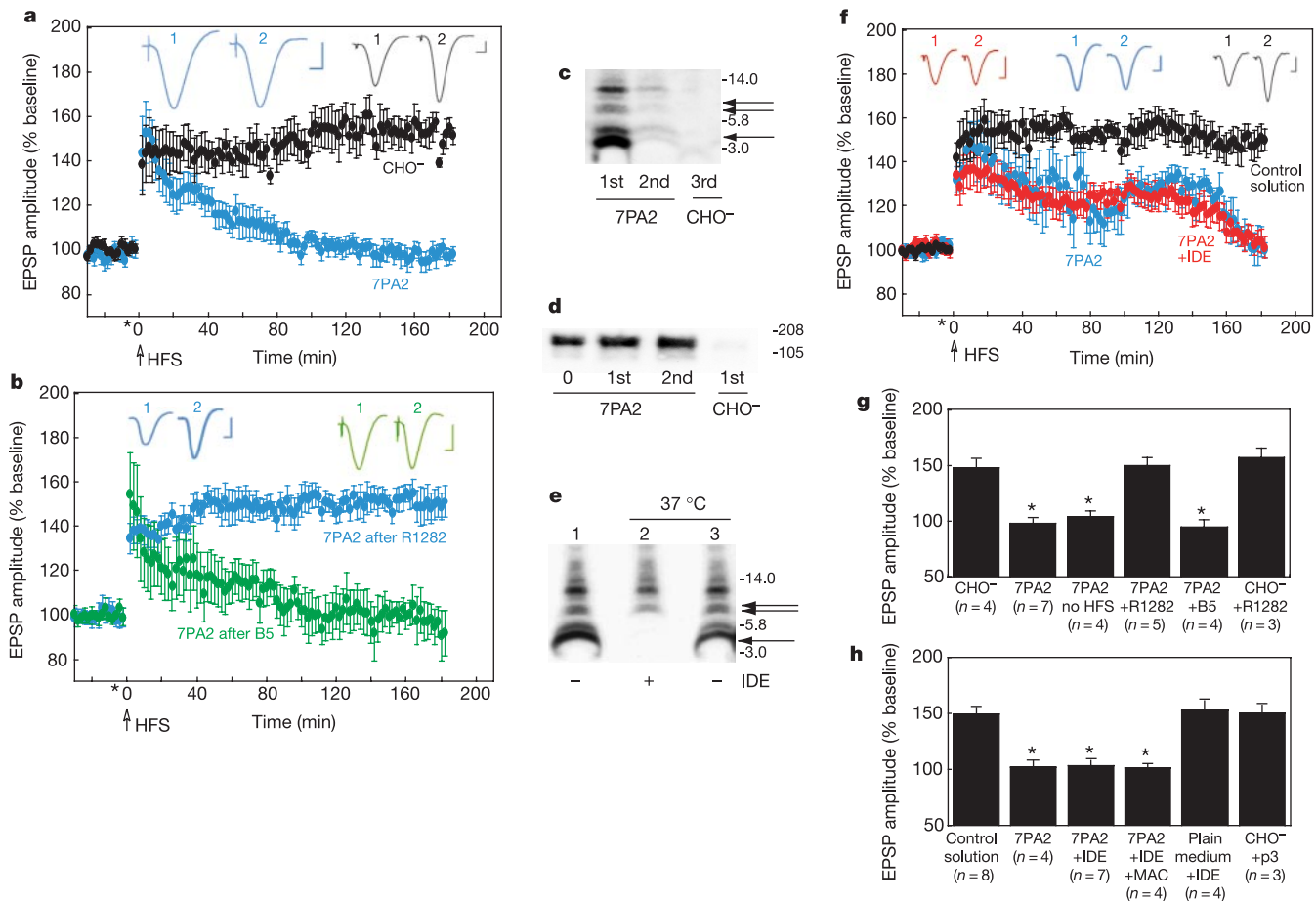


Figure 3 SDS-stable oligomers of human A β block hippocampal LTP *in vivo*. **a**, 7PA2 CM (blue) but not CHO⁻ CM (black, $P < 0.001$) blocked LTP. Insets show typical EPSPs ~ 5 min pre- (1) and ~ 3 h post-HFS (2); calibration bars 5 ms/0.5 mV. **b**, Immunoprecipitation of 7PA2 CM with an antibody to A β (R1282, blue, see **c**) but not to APP_S (B5, green) prevented the block of LTP. **c**, Blots of R1282 precipitates from CM used in **b** show that A β monomer (arrow) and dimer (double arrow) had been efficiently precipitated; compare A β retrieved in 1st precipitate versus in a second precipitation

(2nd). **d**, Blotting of remaining CM of samples in **c** with 22C11 revealed that the R1282 precipitation did not alter APP_S- α levels. **e**, Incubation of 7PA2 CM with IDE caused complete loss of A β monomers (single arrow), whereas A β dimers (double arrow) and trimers were minimally diminished (lanes 1 versus 2). **f**, Treatment of 7PA2 CM with IDE (red) did not prevent the block of LTP. **g**, **h**, Magnitudes of LTP at 3 h post-HFS (% baseline $\pm$ s.e.m.); asterisks $P < 0.05$ compared to CHO⁻ (**g**) or control solution (**h**). n , Number of animals used per treatment.

of LTP maintenance in 4–5 month old²³, 5–7 month old²⁴ or 16 month old²⁵ animals, whereas others find impaired synaptic transmission but no change in LTP at 8–10 months¹² or 12–18 months²⁶. Such studies of transgenic mice lack an advantage of our paradigm, namely the ability to study the effects of biochemically defined assembly forms of naturally produced human A β at physiological levels, in the absence of any confounding effects of APP overexpression.

An attractive therapeutic approach to Alzheimer's disease would be to reduce selectively the levels of potentially synaptotoxic A β oligomers. We show that two chemically distinct, cell-penetrant γ -secretase inhibitors block A β dimer and trimer formation at doses which allow appreciable monomer production. In contrast, anti-

aggregation compounds would need to be both cell-penetrant and capable of blocking very early A β oligomerization; if such molecules did not prevent initial dimerization, they might enhance the levels of potentially synaptotoxic species. Thus, A β -lowering compounds (for example, β - or γ -secretase inhibitors) that prevent intra- and extracellular oligomerization while still allowing significant monomer production would be particularly desirable. Because stable oligomers can potentially arise from a large variety of proteins, both those already implicated in disease²⁷ and those that are not²⁸, the prevention of oligomer formation by reducing monomer concentrations could have wide relevance to the treatment of protein-folding disorders. □

Methods

APP-expressing cells

Chinese hamster ovary cells stably transfected with a complementary DNA coding for APP751 containing the Val717Phe familial Alzheimer's disease mutation (referred to as 7PA2 cells) were cultured in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal bovine serum as described^{8,10}.

Immunoprecipitation/western blot analysis

A β monomers and oligomers were visualized using a highly sensitive immunoprecipitation/western blot protocol that can readily detect as little as 200 pg of naturally secreted human A β ¹⁰. For quantification of band intensity, appropriate film exposures were scanned and the density of bands determined with AlphaEase software (AlphaInnotech).

Subcellular fractionation

Total cellular microsomes were prepared from thirty 10-cm dishes of 7PA2 cells ($\sim 3 \times 10^8$ cells), and fractionated on iodixanol step gradients essentially as described previously¹⁵. The resulting gradient was fractionated into 1-ml aliquots and a 100 $\times$ stock of protease inhibitors and 10% NP40 was added to each fraction to yield 1 $\times$ protease inhibitors and 1% NP40. Fractions were vortexed, and aliquots (50 μ l) removed for western blotting¹⁵. After clearing with a C7-protein A conjugate, samples were examined by immunoprecipitation/western blot¹⁰.

A β ELISA

Enzyme-linked immunosorbent assay (ELISA) for A β _{1–42} was performed as before¹⁰. Full-length APP and APP α were not detected by either assay.

Expression and purification of IDE

Recombinant human IDE (N-terminally tagged with both polyhistidine and haemagglutinin) was expressed in BL21 (DE3) *Escherichia coli* cells and purified using the TALON metal affinity kit (Clontech Laboratories), as described previously²⁹. The specificity of the purified IDE was confirmed by the degradation of ¹²⁵I-A β and ¹²⁵I-insulin and by inhibition with cold insulin or 1,10-phenanthroline¹⁷.

Preparation of conditioned medium for microinjection and electrophysiology

7PA2 and untransfected CHO[−] cells were grown to near confluency and allowed to condition plain DMEM for ~ 16 h. Conditioned medium was cleared of cells (200 $\times$ 10 min, 4 $^{\circ}$ C) and used directly for electrophysiology or else treated to remove A β monomers or both A β monomers and oligomers. To fully deplete monomers but retain oligomers, CM (6 ml) was incubated with recombinant IDE (475 μ g) for 12 h at 37 $^{\circ}$ C. In some experiments, the His-tagged IDE was then removed by metal affinity chromatography²⁹.

Electrophysiology

Experiments were carried out on urethane anaesthetized male adult Wistar rats. Single-pathway recordings of field excitatory postsynaptic potentials (EPSPs) were made from the stratum radiatum in the CA1 area of the hippocampus in response to stimulation of the ipsilateral Schaffer collateral/commissural pathways as described previously¹⁶. Test EPSPs were evoked at a frequency of 0.033 Hz and at a stimulation intensity adjusted to give an EPSP amplitude of 50% of maximum. The HFS protocol for inducing LTP consisted of 10 trains of 20 stimuli, inter-stimulus interval 5 ms (200 Hz), inter-train interval 2 s. The intensity was increased to give an EPSP of 75% of maximum amplitude during the HFS. To inject samples a cannula was implanted in the lateral cerebral ventricle (coordinates: 0.5 mm anterior to bregma and 1.0 mm right of midline) just before electrode implantation¹⁶. Conditioned medium samples (1.5 μ l) were injected over a 2 min period, 10 min before HFS. Control injections comprised superQ water or 0.9% NaCl. LTP was measured as the mean $\pm$ s.e.m. per cent of the baseline field EPSP amplitude recorded over at least a 30-min baseline period. Similar results were obtained when the EPSP slope was measured. Statistical comparisons used paired and unpaired Student's *t*-tests.

Synthesis of γ -secretase inhibitors

The general method of ref. 30 was followed to make MWIII-20, which is carbamic acid, [3-[(1-leucyl-L-leucyl-O-methyl ester)carbonyl](phenylmethyl)amino]-2R-hydroxy-1S-(phenylmethyl)propyl]-1,1-dimethylethyl ester. DAPM (N-[N-3,5-difluorophenyl]-L-alanyl]-L-S-phenylglycine methyl ester) has been reported in patent applications

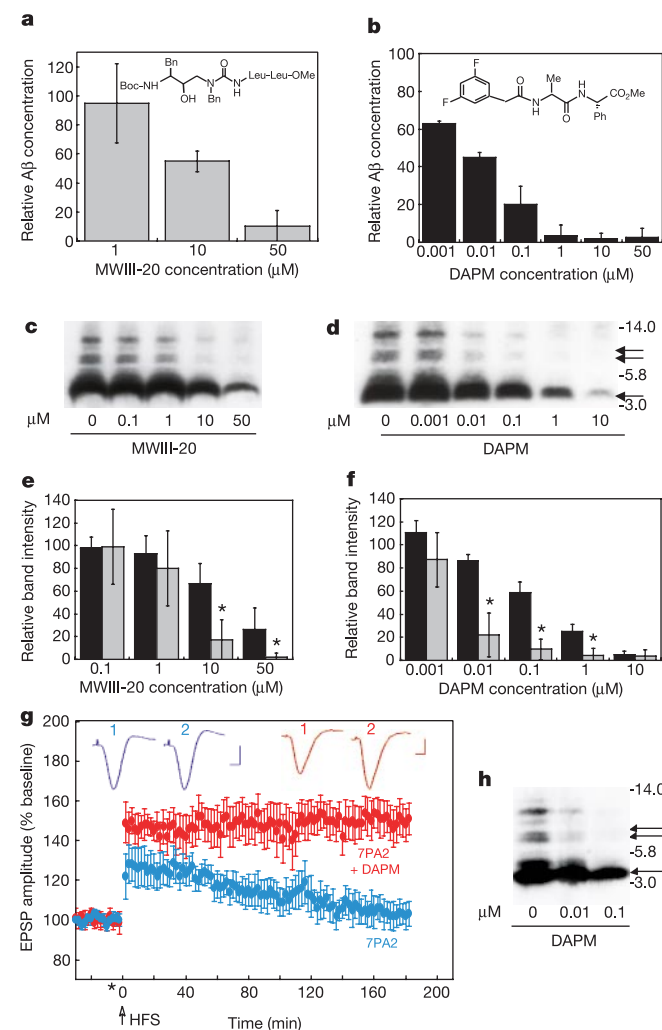


Figure 4 γ -secretase inhibitors block A β oligomer formation at doses that allow substantial monomer production. Inhibition of 7PA2 cellular A β secretion by MWIII-20 (**a**) and DAPM (**b**) was quantified by ELISA. Values (means $\pm$ s.d.) from three independent experiments were normalized to DMSO controls. Insets show inhibitor structures. **c, d**, 7PA2 cells were treated for 6 h with MWIII-20 or DAPM and A β species visualized by immunoprecipitation and western blot. **e, f**, Densitometric quantification of experiments as in **c** and **d**; A β monomer (black bars) and dimer (grey bars) densities normalized to vehicle control. MWIII-20, $n = 8$ and DAPM, $n = 5$. Asterisk indicates that at a given concentration, decrease in dimer is significantly greater ($P < 0.005$) than decrease in monomer. **g**, CM from cells treated with 0.1 μ M DAPM (red, $n = 5$) allowed a robust LTP, whereas media from untreated 7PA2 cells (blue, $n = 4$) blocked LTP (insets as in Fig. 3). **h**, 7PA2 cells were conditioned in plain DMEM with or without DAPM for 6 h; CM was examined by immunoprecipitation and western blot and used for electrophysiology.

WO9822441-A2 and WO9822494-A2 filed by Athena Neurosciences and Eli Lilly. The identity of each compound was confirmed by ¹H-NMR and mass spectrometry.

Received 7 January; accepted 11 February 2002.

- Selkoe, D. J. Alzheimer's disease: genes, proteins and therapies. *Physiol. Rev.* **81**, 742–761 (2001).
- Pike, C. J., Walencewicz, A. J., Glabe, C. G. & Cotman, C. W. *In vitro* aging of β -amyloid protein causes peptide aggregation and neurotoxicity. *Brain Res.* **563**, 311–314 (1991).
- Lorenzo, A. & Yankner, B. A. β -amyloid neurotoxicity requires fibril formation and is inhibited by congo red. *Proc. Natl Acad. Sci. USA* **91**, 12243–12247 (1994).
- Terry, R. D. *et al.* Physical basis of cognitive alterations in Alzheimer's disease: synapse loss is the major correlate of cognitive impairment. *Ann. Neurol.* **30**, 572–580 (1991).
- Dickson, D. W. *et al.* Correlations of synaptic and pathological markers with cognition of the elderly. *Neurobiol. Aging* **16**, 285–298 (1995).
- Lue, L. F. *et al.* Soluble amyloid β peptide concentration as a predictor of synaptic change in Alzheimer's disease. *Am. J. Pathol.* **155**, 853–862 (1999).
- McLean, C. A. *et al.* Soluble pool of A β amyloid as a determinant of severity of neurodegeneration in Alzheimer's disease. *Ann. Neurol.* **46**, 860–866 (1999).
- Podlisny, M. B. *et al.* Aggregation of secreted amyloid β -protein into sodium dodecyl sulfate-stable oligomers in cell culture. *J. Biol. Chem.* **270**, 9564–9570 (1995).
- Morishima-Kawashima, M. & Ihara, Y. The presence of amyloid β -protein in the detergent-insoluble membrane compartment of human neuroblastoma cells. *Biochemistry* **37**, 15247–15253 (1998).
- Walsh, D. M., Tseng, B. P., Rydel, R. E., Podlisny, M. B. & Selkoe, D. J. Detection of intracellular oligomers of amyloid β -protein in cells derived from human brain. *Biochemistry* **39**, 10831–10839 (2000).
- Xia, W. M. *et al.* Enhanced production and oligomerization of the 42-residue amyloid β -protein by Chinese hamster ovary cells stably expressing mutant presenilins. *J. Biol. Chem.* **272**, 7977–7982 (1997).
- Hsia, A. Y. *et al.* Plaque-independent disruption of neural circuits in Alzheimer's disease mouse model. *Proc. Natl Acad. Sci. USA* **96**, 3228–3233 (1999).
- Mucke, L. *et al.* High-level neuronal expression of A β 1–42 in wild-type human amyloid protein precursor transgenic mice: synaptotoxicity without plaque formation. *J. Neurosci.* **20**, 4050–4058 (2000).
- Hartley, D. M. *et al.* Protofibrillar intermediates of amyloid β -protein induce acute electrophysiological changes and progressive neurotoxicity in cortical neurons. *J. Neurosci.* **19**, 8876–8884 (1999).
- Xia, W. *et al.* Presenilin complexes with the C-terminal fragments of amyloid precursor protein at the sites of amyloid β -protein generation. *Proc. Natl Acad. Sci. USA* **97**, 9299–9304 (2000).
- Kim, J. H., Anwyl, R., Suh, Y. H., Djamgoz, M. B. & Rowan, M. J. Use-dependent effects of amyloidogenic fragments of β -amyloid precursor protein on synaptic plasticity in rat hippocampus *in vivo*. *J. Neurosci.* **21**, 1327–1333 (2001).
- Qiu, W. Q. *et al.* Insulin-degrading enzyme regulates extracellular levels of amyloid β -protein by degradation. *J. Biol. Chem.* **273**, 32730–32738 (1998).
- Lambert, M. P. *et al.* Diffusible, nonfibrillar ligands derived from A β _{1–42} are potent central nervous system neurotoxins. *Proc. Natl Acad. Sci. USA* **95**, 6448–6453 (1998).
- Walsh, D. M. *et al.* Amyloid β -protein fibrillogenesis: structure and biological activity of protofibrillar intermediates. *J. Biol. Chem.* **274**, 25945–25952 (1999).
- Chui, D.-H. *et al.* Transgenic mice with Alzheimer presenilin 1 mutations show accelerated neurodegeneration without amyloid plaque formation. *Nature Med.* **5**, 560–564 (1999).
- Chen, G. *et al.* A learning deficit related to age and β -amyloid plaques in a mouse model of Alzheimer's disease. *Nature* **408**, 975–979 (2000).
- Lewis, J. *et al.* Enhanced neurofibrillary degeneration in transgenic mice expressing mutant tau and APP. *Science* **293**, 1487–1491 (2001).
- Larson, J., Lynch, G., Games, D. & Seubert, P. Alterations in synaptic transmission and long-term potentiation in hippocampal slices for young and aged PDAPP mice. *Brain Res.* **840**, 23–35 (1999).
- Moehrs, D. *et al.* Early phenotypic changes in transgenic mice that overexpress different mutants of amyloid precursor protein in brain. *J. Biol. Chem.* **274**, 6483–6492 (1999).
- Chapman, P. F. *et al.* Impaired synaptic plasticity and learning in aged amyloid precursor protein transgenic mice. *Nature Neurosci.* **2**, 271–276 (1999).
- Fitzhohn, S. M. *et al.* Age-related impairment of synaptic transmission but normal long-term potentiation in transgenic mice that overexpress the human APP695SWE mutant form of amyloid precursor protein. *J. Neurosci.* **21**, 4691–4698 (2001).
- Rochet, J. C. & Lansbury, P. T. Jr Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* **10**, 60–68 (2000).
- Chiti, F. *et al.* Designing conditions for *in vitro* formation of amyloid protofibrils and fibrils. *Proc. Natl Acad. Sci. USA* **96**, 3590–3590 (1999).
- Chesneau, V. & Rosner, M. R. Functional human insulin-degrading enzyme can be expressed in bacteria. *Protein Expr. Purif.* **19**, 91–98 (2000).
- Getman, D. P. *et al.* Discovery of a novel class of potent HIV-1 protease inhibitors containing the (R)-(hydroxyethyl)urea isostere. *J. Med. Chem.* **36**, 288–291 (1993).

Acknowledgements

We thank M. Rosner and V. Chesneau for the gift of the pProExH6HA IDE expression vector, B. Zheng for ELISA analysis, S. Mansourian for assistance in the preparation of illustrations and W. T. Kimberly, W. P. Esler and D. M. Hartley for discussions. Supported by NIH grants (to D.J.S. and M.S.W.) and by Enterprise Ireland and the Health Research Board Ireland (M.R. and R.A.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.J.S. (e-mail: dselkoe@rics.bwh.harvard.edu).

A global disorder of imprinting in the human female germ line

Hannah Judson, Bruce E. Hayward, Eamonn Sheridan & David T. Bonthron

University of Leeds, Molecular Medicine Unit, St. James's University Hospital, Leeds LS9 7TF, UK

Imprinted genes are expressed differently depending on whether they are carried by a chromosome of maternal or paternal origin. Correct imprinting is established by germline-specific modifications; failure of this process underlies several inherited human syndromes^{1–5}. All these imprinting control defects are *cis*-acting, disrupting establishment or maintenance of allele-specific epigenetic modifications across one contiguous segment of the genome. In contrast, we report here an inherited global imprinting defect. This recessive maternal-effect mutation disrupts the specification of imprints at multiple, non-contiguous loci, with the result that genes normally carrying a maternal methylation imprint assume a paternal epigenetic pattern on the maternal allele. The resulting conception is phenotypically indistinguishable from an androgenetic complete hydatidiform mole⁶, in which abnormal extra-embryonic tissue proliferates while development of the embryo is absent or nearly so. This disorder offers a genetic route to the identification of *trans*-acting oocyte factors that mediate maternal imprint establishment.

Although normally sporadic, complete hydatidiform mole (CHM) is occasionally familial, with affected women repeatedly having pregnancies of this type. These repetitive CHMs are not androgenetic but biparental (BiCHM)^{7–9}. By analogy to disorders like Prader–Willi syndrome (which can result from sporadic uniparental disomy or from familial imprinting control mutations), we considered that BiCHM might arise from a global inherited failure of maternal imprinting.

We studied the sixth molar pregnancy of the index case in a BiCHM family with complex consanguinity, originating from the Mirpur region of Pakistan. We demonstrated biparental origin of the BiCHM DNA using markers on six autosomes.

Imprinted genes are associated with differentially methylated regions (DMRs), either 'primary' (established during gametogenesis) or 'secondary' (established later in embryogenesis). We used bisulphite sequencing¹⁰ to compare methylation in the BiCHM and suitable controls, including uniparental DNAs and first-trimester chorionic villus samples, which like CHMs, are of trophoblastic origin.

The Beckwith–Wiedemann region of 11p15 contains two putative primary imprint control regions, at *H19* and *KCNQ1OT1*, ~500 kilobases (500 kb) apart. The DMR ~2-kb upstream of *H19* normally shows paternal-specific germline methylation¹¹, and is therefore an important control locus (Fig. 1a). Parthenogenetic (Pg) and androgenetic (Ag) control DNAs were respectively completely unmethylated and completely methylated at all CpG dinucleotides, as expected. The BiCHM DNA shows a differentially methylated pattern, like that of normal controls. Cloned polymerase chain reaction (PCR) products from BiCHM were either almost completely methylated or completely unmethylated, as expected for paternal or maternal alleles, respectively. This maintenance of normal *H19* differential methylation in the BiCHM is as predicted, if only imprinting in the female germ line is affected.

At loci with a maternal methylation imprint (Fig. 1b–e), a very different pattern is seen. The *KCNQ1OT1* primary DMR^{12,13} becomes methylated during oogenesis¹⁴. As expected, our normal control DNAs are uniformly haplo-methylated (C and T bands of similar intensity at each original CpG position), and the partheno-

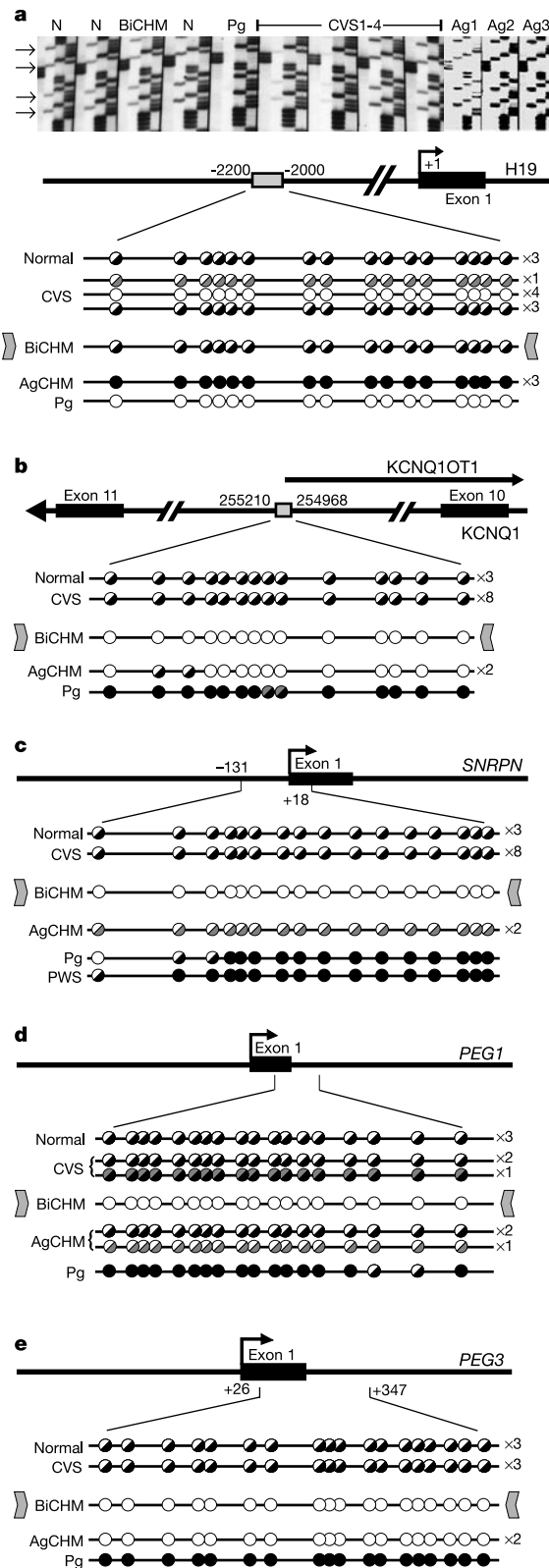


Figure 1 Bisulphite sequencing of DMRs in imprinted genes. Circles represent positions and methylation of individual CpGs, as follows: filled, methylated; open, unmethylated; half-filled black/white, haplo-methylated; black/grey, predominantly methylated; grey/white, predominantly unmethylated. Identical results from multiple controls are collated, numbers indicated to the right. **a**, *H19*. The first and last CpG are numbered relative to the transcriptional start site. Arrows indicate differentially methylated C residues on the gel. N, adult control DNA. Lanes: ACGT, left to right. **b**, *KCNQ1OT1*. Numbering refers to accession AJ006345. **c**, *SNRPN*. Numbering relates to the first nucleotide of exon 1. PWS, Prader–Willi syndrome. **d**, *PEG1*. **e**, *ZIM2/PEG3*. CVS, chorionic villus sample.

genetic sample fully methylated. In contrast, the BiCHM DNA is completely unmethylated, its maternal *KCNQ1OT1* allele thus having a paternal epigenotype.

The 5' DMR of *SNRPN* (15q) behaved similarly. In the mouse, this is a primary imprint¹⁵, but in humans may only become established in early post-zygotic development¹⁶. In the BiCHM, this DMR was completely unmethylated (paternal epigenotype), whereas the parthenogenetic and Prader–Willi controls had the opposite epigenotype (almost all CpGs completely methylated). Chorionic villus samples and other normal controls were uniformly haplo-methylated (Fig. 1c). AgCHM were hypomethylated compared to normal DNA, but unlike the BiCHM did show faint bands indicating some (presumably secondary) CpG methylation (see also Supplementary Information).

PEG1 (7q32) and *ZIM2/PEG3* (19q13.4)^{17,18} both have maternally methylated DMRs. It is not known if these are primary imprints, although the demethylated paternal *PEG1* epigenotype is established during spermatogenesis¹¹. In the BiCHM, these DMRs are both completely unmethylated (paternal epigenotype on both alleles). At *ZIM2/PEG3*, the controls appear as predicted, the normal DNAs being haplo-methylated, the PgDNA completely methylated, and the AgCHM, like the BiCHM, unmethylated (Fig. 1e). However, at *PEG1*, whilst the normal and parthenogenetic samples are respectively haplo-methylated and completely methylated (as expected) the AgCHM DNAs show a variable degree of incomplete methylation (Fig. 1d).

To test whether the BiCHM methylation abnormalities truly reflect a defect of maternal gametic imprinting, rather than being secondary to the molar phenotype, we examined a complex locus, *GNAS1*, that has multiple imprinted transcripts and at least three separate DMRs^{19–22} (Fig. 2). In murine *Gnas*, the exon 1A DMR is a primary imprint, whereas the upstream DMRs only become established during the blastocyst stage²². Likewise, *GNAS1* imprinting mutations that cause type Ib pseudohypoparathyroidism (PHP-Ib) always alter exon 1A methylation, with the other DMRs only sometimes affected⁵. Therefore, a maternal germline imprinting defect should involve failure to methylate the maternal allele of exon 1A. The maternal NESP55, XL α s, and antisense promoter DMRs (all 35–50 kb upstream) should then secondarily assume a paternal epigenotype, becoming respectively methylated, unmethylated, and unmethylated. This prediction was almost completely fulfilled. At exon 1A, the parthenogenetic control, as expected, is completely methylated, whereas the BiCHM is completely unmethylated, indicating failure to establish the maternal primary imprint. There is some variability in methylation in control samples; two of ten chorionic villus sample DNAs are hypomethylated, and one of three AgCHM appears partially methylated, suggesting that some secondary methylation must have appeared at this locus. Nonetheless, the unmethylated status of the BiCHM is as predicted, and that this represents a true germline defect is supported by analysis of the other DMRs at this locus.

The NESP55 DMR becomes methylated on the paternal allele in the blastocyst stage, possibly secondary to antisense transcription^{21,22}. BiCHM and AgCHM are both completely methylated at this DMR (paternal epigenotype on both alleles). All other controls show the expected methylation patterns. Thus, the postzygotic mechanism that sets up the secondary paternal NESP55 imprint remains operative in the BiCHM, but in the absence of a maternal gametic imprint at 1A this yields a paternal methylation pattern on both, rather than one, NESP55 alleles.

We also examined two regions ~3 kb apart, within a large (5-kb) CpG island spanning the antisense promoter and XL α s exon. At the antisense DMR, the BiCHM again shows a paternal epigenotype (this time unmethylated) on both alleles. This lack of methylation is distinctive, even though both AgCHMs show a minor degree of secondary methylation at this locus.

The maternal XL α s allele becomes methylated during the blas-

tocyst stage²². Here we initially saw no sign of abnormal methylation in the BiCHM, the DNA appearing haplo-methylated. Similar partial methylation at this DMR, independent of correct maternal methylation at exon 1A, has been seen with *cis*-acting *GNAS1* imprinting defects that cause PHP-Ib; several such families have an abnormal (paternal) methylation pattern at exon 1A and NESP55, whereas the XL α s DMR appears unaffected⁵. Cloning of the BiCHM bisulphite-PCR products, however, revealed a disordered pattern of partial methylation scattered irregularly across the clones, rather than the normal grouping into completely methylated and completely unmethylated clones (see Supplementary Information). A similar analysis has not been reported for the PHP-Ib mutations. Thus, despite the appearance of some methylation at the XL α s DMR, the overall evidence from the four DMRs argues compellingly for a *GNAS1* imprinting defect in BiCHM, very similar to that resulting from some maternally transmitted *cis*-acting imprinting mutations.

The contrasting behaviour of *H19* and *GNAS1*-NESP55 in the BiCHM is noteworthy. Although both DMRs are normally paternally methylated, for *H19* this is primary, and therefore unaffected by an oocyte defect. At NESP55, paternal methylation is secondary to lack of a maternal imprint at 1A, and hence occurs on both alleles in the BiCHM. This difference suggests that the BiCHM defect is not a generalized failure of methylation maintenance, but reflects specific events in the female germ line. Also consistent with this conclusion was the observation of a normal methylated status in the BiCHM at eight CpGs in an intragenic (non-CpG island) region of

the non-imprinted *KHK* gene (not shown). Other evidence argues that the BiCHM methylation abnormalities reflect a specific imprinting defect, rather than changes peculiar to trophoblast derivatives. First, we see neither random nor generalized hypo- or hyper-methylation; instead, at each DMR, the direction of the BiCHM methylation abnormality is specifically as predicted for a maternal germline defect. Second, despite some minor inter-sample variation, chorionic villus sample DNAs (which, like the CHMs, are first trimester trophoblast derivatives) typically had normal differential methylation, and never had a 'paternal-only' epigenotype resembling that of the AgCHM and BiCHM.

Cis-acting mutations that disrupt imprinting at individual loci^{1-5,12,13,23-24} show sex-dependent dominant (vertical) transmission. In contrast, BiCHM is a pure maternal-effect defect; affected women may have molar pregnancies with different partners⁹, but are otherwise healthy. Its presumed autosomal recessive inheritance pattern implies a *trans*-acting molecular defect, consistent with the involvement of multiple dispersed imprinted loci. Of all imprinted loci examined, only *H19* (as predicted) showed a normal differentially methylated pattern in the BiCHM. Because most gametic imprints are imposed in the female rather than the male germ line²⁵, the great majority of all imprinted loci are probably affected by this genetic defect. A recessive *Dnmt3L* mutation, although also conferring male sterility, prevents specification of maternal imprints in the mouse germ line²⁶. In the family studied here, lack of homozygosity for the corresponding human locus, as well as for a previously suggested 19q BiCHM locus⁷ (see

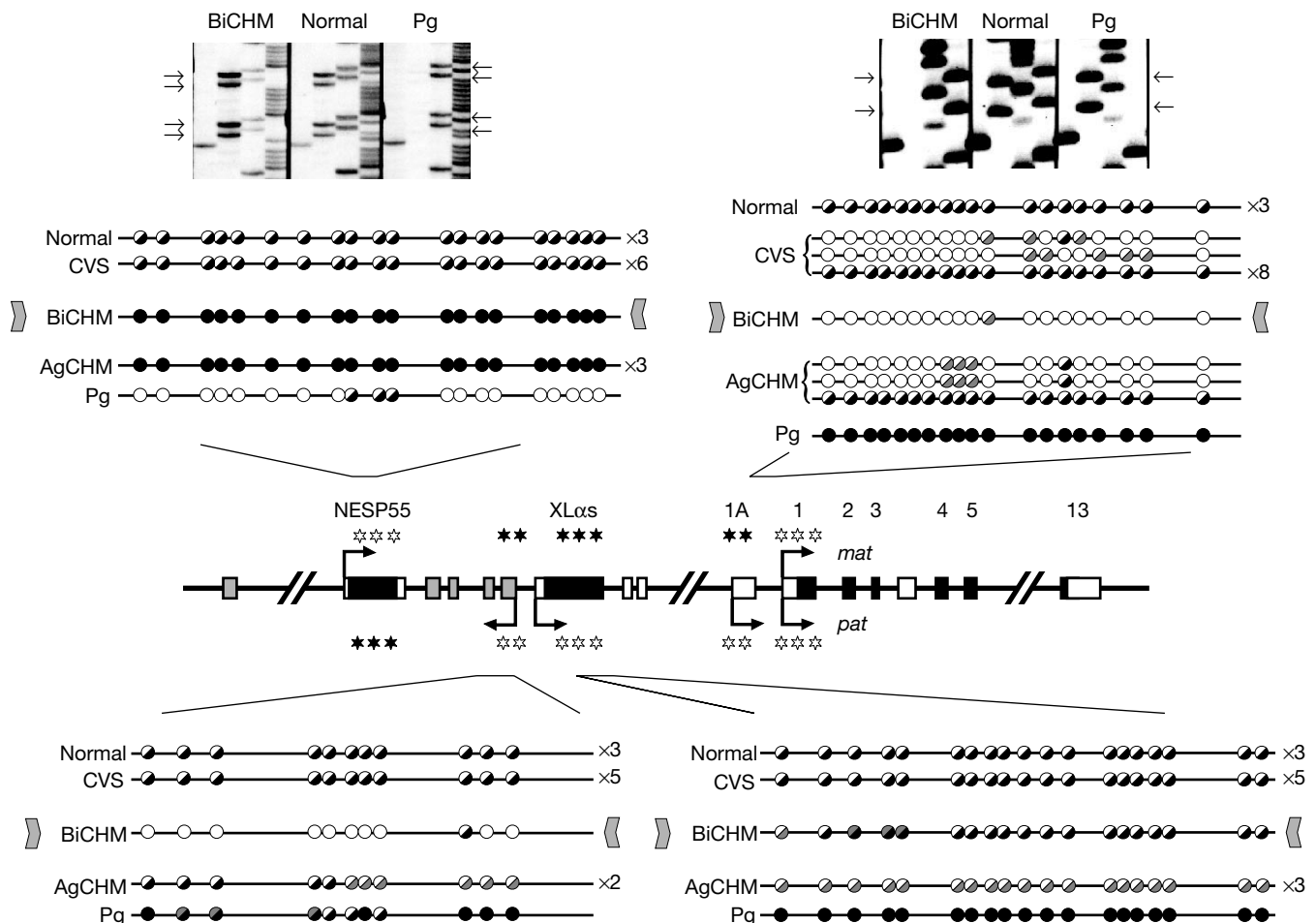


Figure 2 Bisulphite analysis of *GNAS1*. At the top, the contrasting methylation status at the 1A and NESP55 DMRs is illustrated; arrows on the gels indicate differentially methylated residues. The remainder of the figure summarizes data for the whole locus (for symbols see Fig. 1). In the schematic map of *GNAS1*, coding regions are black, antisense

exons grey. All alternative first exons on the sense strand splice onto exon 2 (refs 5, 21). Arrows above and below the line indicate maternal and paternal transcription, respectively. Normally methylated or unmethylated status of CpG islands is shown by filled or open stars, respectively.

Supplementary Information) makes involvement of either of these loci unlikely. However, the BiCHM defect should eventually be identifiable through autozygosity mapping. □

Methods

Detailed methods are available in Supplementary Information.

DNA samples

BiCHM DNA was extracted from a short-term culture of the evacuation products from the sixth pregnancy of the index case. Four of her first five conceptions had previously been histologically confirmed as CHM, and demonstrated to be biparental using archival pathological material. Parthenogenetic DNA was previously described²⁷. Adult control blood DNAs were from the index case, her husband, and an unrelated individual. Fluorescent PCR analysis of markers D1S2691, D5S495, D10S189, D13S1293, D17S946, D19S210 and D19S413 was performed by standard methods on DNA from the cultured BiCHM and from the index case and her husband; all these markers were fully informative for demonstrating both maternal and paternal allelic contributions to the mole.

Bisulphite-PCR analysis of DNA methylation

The protocol was adapted from previously described methods^{10,24}. Briefly, genomic DNA was denatured and bisulphite-treated to convert unmethylated cytosines to thymines. PCR products encompassing the DMRs of each imprinted locus were then generated. Only one strand was amplified at each locus. Products were analysed by direct sequencing, because at some loci the two modified alleles clone with different efficiencies. Cloning was therefore used only to assess the allelic separation of C and T at haplo-methylated loci (see text).

Received 18 September 2001; accepted 21 January 2002.

1. Sutcliffe, J. S. *et al.* Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nature Genet.* **8**, 52–58 (1994).
2. Buiting, K. *et al.* Inherited microdeletions in the Angelman and Prader-Willi syndromes define an imprinting centre on human chromosome 15. *Nature Genet.* **9**, 395–400 (1995).
3. Reik, W. *et al.* Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by altered imprinting pattern in the IGF2-H19 domain. *Hum. Mol. Genet.* **4**, 2379–2385 (1995).
4. Gardner, R. J. *et al.* An imprinted locus associated with transient neonatal diabetes mellitus. *Hum. Mol. Genet.* **9**, 589–596 (2000).
5. Liu, J. *et al.* A GNAS1 imprinting defect in pseudohypoparathyroidism type IB. *J. Clin. Invest.* **106**, 1167–1174 (2000).
6. Kajii, T. & Ohama, K. Androgenetic origin of hydatidiform mole. *Nature* **268**, 633–634 (1977).
7. Mogilabey, Y. B. *et al.* Genetic mapping of a maternal locus responsible for familial hydatidiform moles. *Hum. Mol. Genet.* **8**, 667–671 (1999).
8. Helwani, M. N. *et al.* A familial case of recurrent hydatidiform molar pregnancies with biparental genomic contribution. *Hum. Genet.* **105**, 112–115 (1999).
9. Fisher, R. A., Khatoun, R., Paradinas, F. J., Roberts, A. P. & Newlands, E. S. Repetitive complete hydatidiform mole can be biparental in origin and either male or female. *Hum. Reprod.* **15**, 594–598 (2000).
10. Clark, S. J., Harrison, J., Paul, C. L. & Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997 (1994).
11. Kerjean, A. *et al.* Establishment of the paternal methylation imprint of the human H19 and MEST/PEG1 genes during spermatogenesis. *Hum. Mol. Genet.* **9**, 2183–2187 (2000).
12. Lee, M. P. *et al.* Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proc. Natl Acad. Sci. USA* **96**, 5203–5208 (1999).
13. Ohta, T. *et al.* Imprinting-mutation mechanisms in Prader-Willi syndrome. *Am. J. Hum. Genet.* **64**, 397–413 (1999).
14. Engemann, S. *et al.* Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. *Hum. Mol. Genet.* **9**, 2691–2706 (2000).
15. Shemer, R., Birger, Y., Riggs, A. D. & Razin, A. Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. *Proc. Natl Acad. Sci. USA* **94**, 10267–10272 (1997).
16. El-Maarri, O. *et al.* Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nature Genet.* **27**, 341–344 (2001).
17. Kaneko-Ishino, T. *et al.* Peg1/Mest imprinted gene on chromosome 6 identified by cDNA subtraction hybridization. *Nature Genet.* **11**, 52–59 (1995).
18. Murphy, S. K., Wylie, A. A. & Jirtle, R. L. Imprinting of PEG3, the human homologue of a mouse gene involved in nurturing behavior. *Genomics* **71**, 110–117 (2001).
19. Hayward, B. E. *et al.* The human GNAS1 gene is imprinted and encodes distinct paternally and biallelically expressed G proteins. *Proc. Natl Acad. Sci. USA* **95**, 10038–10043 (1998).
20. Hayward, B. E., Moran, V., Strain, L. & Bonthron, D. T. Bidirectional imprinting of a single gene: GNAS1 encodes maternally, paternally, and biallelically derived proteins. *Proc. Natl Acad. Sci. USA* **95**, 15475–15480 (1998).
21. Hayward, B. E. & Bonthron, D. T. An imprinted antisense transcript at the human GNAS1 locus. *Hum. Mol. Genet.* **9**, 835–841 (2000).
22. Liu, J., Yu, S., Litman, D., Chen, W. & Weinstein, L. S. Identification of a methylation imprint mark within the mouse Gnas locus. *Mol. Cell Biol.* **20**, 5808–5817 (2000).
23. Smilnich, N. J. *et al.* A maternally methylated CpG island in KVLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc. Natl Acad. Sci. USA* **96**, 8064–8069 (1999).
24. Kamiya, M. *et al.* The cell cycle control gene ZAC/PLAGL1 is imprinted—a strong candidate gene for transient neonatal diabetes. *Hum. Mol. Genet.* **9**, 453–460 (2000).
25. Reik, W. & Walter, J. Evolution of imprinting mechanisms: the battle of the sexes begins in the zygote. *Nature Genet.* **27**, 255–256 (2001).

26. Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B. & Bestor, T. H. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**, 2536–2539 (2001); advance online publication, 29 November 2001 (DOI 10.1126/Science.1065848).
27. Strain, L., Warner, J. P., Johnston, T. & Bonthron, D. T. A human parthenogenetic chimera. *Nature Genet.* **11**, 164–169 (1995).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank R. Fisher for supplying androgenetic CHM DNAs, and G. Taylor for Prader-Willi and chorionic villus sample DNA samples. This work was supported by the Wellcome Trust.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to D.T.B. (e-mail: d.t.bonthron@leeds.ac.uk).

Bone marrow cells adopt the phenotype of other cells by spontaneous cell fusion

Naohiro Terada*†, Takashi Hamazaki*, Masahiro Oka*, Masanori Hoki*, Diana M. Mastalerz*, Yuka Nakano‡, Edwin M. Meyer‡, Laurence Morel*, Bryon E. Petersen*† & Edward W. Scott†§

* Department of Pathology, † Program in Stem Cell Biology, Shands Cancer Center, ‡ Department of Pharmacology, § Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, Florida 32610, USA

Recent studies have demonstrated that transplanted bone marrow cells can turn into unexpected lineages including myocytes, hepatocytes, neurons and many others¹. A potential problem, however, is that reports discussing such 'transdifferentiation' *in vivo* tend to conclude donor origin of transdifferentiated cells on the basis of the existence of donor-specific genes such as Y-chromosome markers¹. Here we demonstrate that mouse bone marrow cells can fuse spontaneously with embryonic stem cells in culture *in vitro* that contains interleukin-3. Moreover, spontaneously fused bone marrow cells can subsequently adopt the phenotype of the recipient cells, which, without detailed genetic analysis, might be interpreted as 'dedifferentiation' or transdifferentiation.

Recent progress in stem cell research indicates that certain mammalian cells, even from adults, maintain a high degree of plasticity for multilineage cell differentiation. The transferred nuclei from adult cells could be reprogrammed by a factor or factors in the cytoplasm of oocytes, showing the same potential for normal animal development as early embryonic nuclei². More recently, neural stem cells were demonstrated to differentiate into virtually every cell type when they were injected into blastocysts *in vivo* or cultured *in vitro* with differentiating embryonic stem cells³. This indicated that the extracellular factor(s) of blastocysts or embryonic stem cells, or cell-cell interaction of neural stem cells with such embryonic cells, might be sufficient for reprogramming adult cells into a more pluripotent status. To this end, we attempted to establish a culture of pluripotent stem cells *in vitro* from adult cells (bone marrow cells) by nurturing them with embryonic stem cells. Bone marrow contains haematopoietic stem cells producing

all the blood cells, and mesenchymal stem cells⁴. Moreover, transplantation of adult bone marrow cells has generated unexpected phenotypes *in vivo*, including muscle cells⁵, liver cells^{6–8}, brain cells^{9,10} and others¹¹. These results indicate that adult bone marrow might contain pluripotent stem cells, or stem cells that could become pluripotent under appropriate conditions.

Mouse bone marrow cells were obtained from the femurs of 8–9-week-old female transgenic mice expressing green fluorescent protein (GFP) and the puromycin-resistance gene¹². Bone marrow mononuclear cells (1×10^6) were mixed with mouse male embryonic stem cells (R1, 1×10^4 cells) in embryonic stem culture medium containing leukaemia inhibitory factor (LIF)¹³, and plated on a gelatin-coated culture dish. To support growth of haematopoietic cells, interleukin (IL)-3 was added to the culture for the initial 7 days. Floating cells were sequentially gathered on days 3 and 7, and transferred to a new gelatin-coated plate each time with the same medium. To remove embryonic stem cells, puromycin was added to the culture after day 7. Within 3 weeks, we found multiple GFP-positive clones morphologically similar to embryonic stem cells (growing as a flattened colony with tight cell–cell conjunction) (Fig. 1a). These clones proliferated at a similar rate to control embryonic stem cells for more than 6 months. GFP expression and resistance to puromycin demonstrated that the GFP-positive embryonic stem-like cells were derived from bone marrow cells.

The bone marrow-derived, embryonic stem-like (BMESL) cells were then subjected to the *in vitro* differentiation protocol of embryonic stem cells (Fig. 1b)^{14,15}. On removal of LIF, BMESL cells formed a sphere of aggregated cells in a floating culture (Fig. 1b, day 3), and differentiated into various morphologies including beating cardiac myocytes after the aggregated cells were attached to a cell culture dish (Fig. 1b, day 15). Gene expression was examined by polymerase chain reaction with reverse transcription (RT–PCR) (Fig. 1c). Oct3/4 and UTF1, markers of undifferentiated

embryonic stem cells^{16,17}, were highly expressed in BMESL cells, and expression of them decreased when BMESL cells differentiated. In contrast, expression of genes indicating mesodermal differentiation (brachyury and collagen II), endodermal differentiation (α -feto-protein and albumin) and ectodermal differentiation (neurofilament H, dopamine β hydroxylase and PTX3) were induced after removal of LIF. Additionally, when BMESL cells were cultured on PA6 cells¹⁸, cells were predominantly differentiated into morphologically neuron-like cells expressing tyrosine hydroxylase (data not shown). Further, BMESL cells were able to form teratomas when injected into immunocompromised NOD/SCID mice (Fig. 1d).

The data suggested the establishment of embryonic stem-like pluripotent stem cells from adult bone marrow cells by nurturing them with embryonic stem cells. However, genetic analyses of the BMESL cells led to an alternative, unexpected conclusion. DNA ploidy (the number of DNA copies) was initially examined by fluorescence-activated cell sorting (FACS analysis) after staining DNA with propidium iodide. All 13 BMESL clones examined demonstrated over-diploid DNA content. Eleven clones showed approximately tetraploid ($4n$) DNA content, and two showed approximately hexaploid ($6n$) DNA content (Fig. 2a). When the karyotype was examined in two tetraploid lines, each had XXXY sex chromosomes, and their modal number of chromosomes was between 78 and 79 (Fig. 2b). Further, BMESL cells were shown to have a hybrid genotype of embryonic stem cells and bone marrow cells by PCR amplification of multiple microsatellites, which were polymorphic between the bone marrow cell genome (mostly C57BL/6) and embryonic stem cell genome (129/Sv) (Fig. 2c). These data indicated that BMESL cells were generated by cell fusion of embryonic stem and bone marrow cells. Blastocyst injection of BMESL cells has so far failed to generate chimaeric mice with a green fluorescent phenotype, which may be explained by their tetraploid genotype¹⁹. It should be noted that dominance of the pluripotent

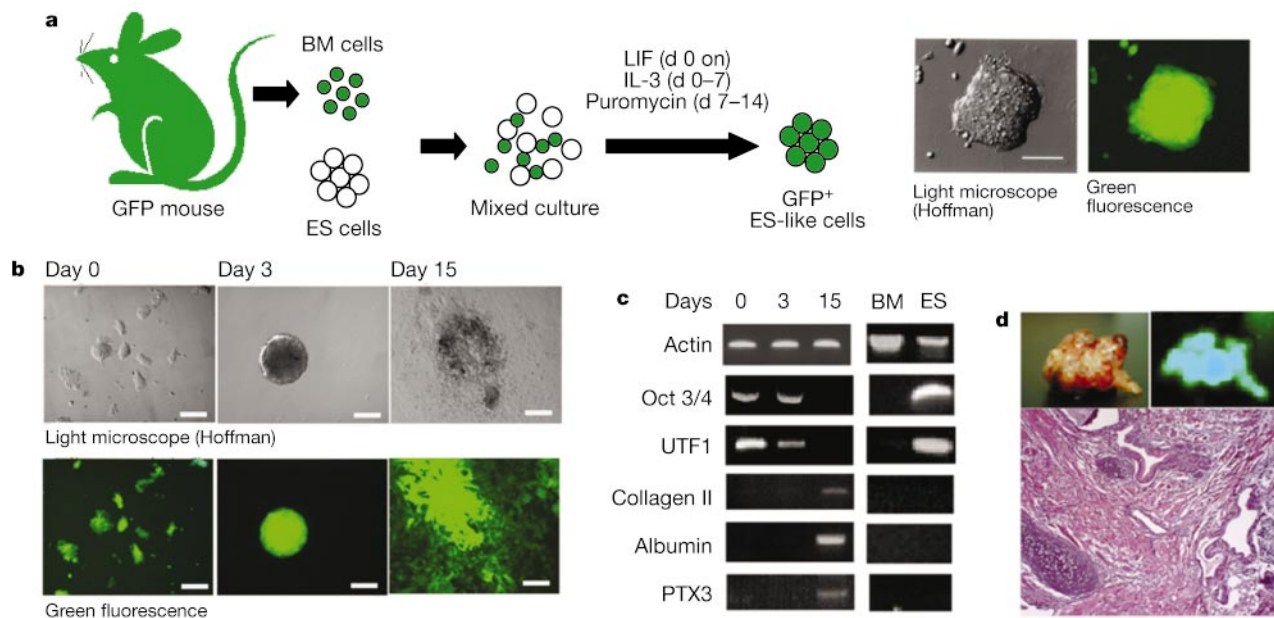


Figure 1 GFP⁺ embryonic stem-like cells derived from mixed culture of GFP⁺ bone marrow cells and GFP[−] embryonic stem cells. **a**, Mouse bone marrow (BM) cells from female transgenic mice expressing GFP (see Methods) were cocultured with male embryonic stem (ES) cells (R1) in the embryonic stem culture medium containing 1,000 U ml^{−1} recombinant mouse LIF in the presence of IL-3 (1 ng ml^{−1}, days 0–7) and puromycin (5 μ g ml^{−1}, days 7–14) for the indicated times. GFP⁺ clones morphologically similar to embryonic stem cells were found in the culture within 3 weeks. Scale bars, 100 μ m. **b**, *In vitro* differentiation of BMESL cells (clone number 1). Scale bars, 200 μ m. **c**, Total RNA was extracted from differentiating BMESL cells (days 0, 3, 15), bone marrow

cells and embryonic stem cells. Complementary DNA was synthesized from 2 μ g total RNA. For each gene, the DNA primers were derived from different exons to ensure that the PCR product represents the specific mRNA species and not genomic DNA. The DNA sequence of the primers is given in the Supplementary Information. **d**, BMESL cells (1×10^5 cells) were injected into the spleen of a NOD/SCID mouse. In four weeks, a tumour (about 2.5 cm diameter) was formed on spleen (top left), which was GFP⁺ under a hand ultraviolet irradiator (top right). The tumour contained a variety of cell types including chondrocytes, striated muscle cells and gland-like structures shown by haematoxylin–eosin staining (light microscope, $\times 100$).

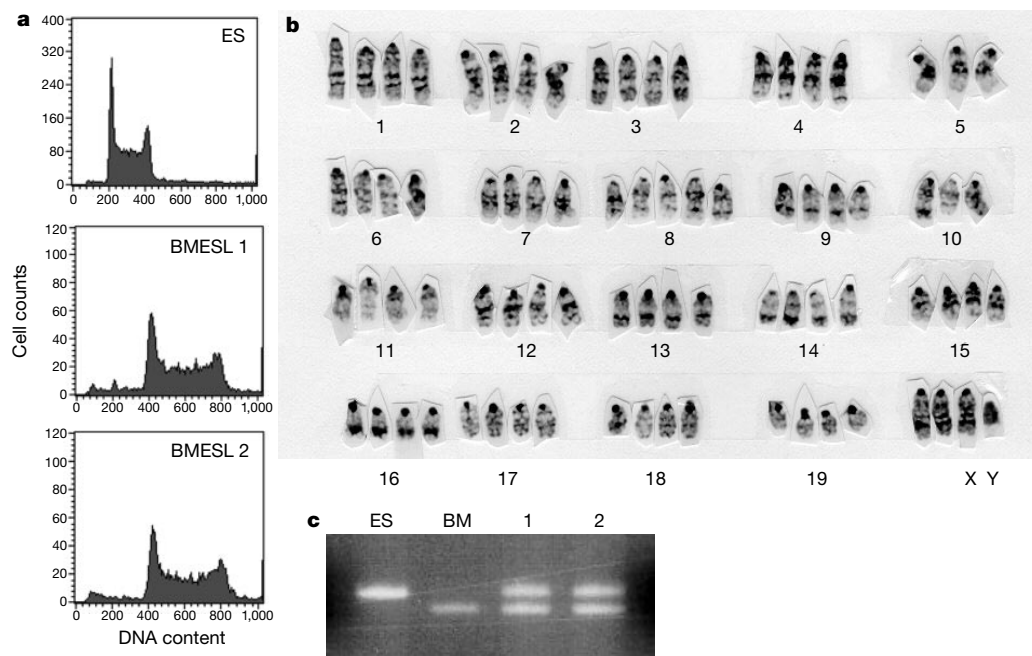


Figure 2 Genetic analysis of BMESL cells. **a**, DNA ploidy. Parental embryonic stem (ES) cells and BMESL cells (clone numbers 1 and 2, 4 weeks after cloning) were stained with propidium iodide and subjected to FACS analysis. **b**, Karyotype (BMESL clone 1). **c**, DNA polymorphism. Genomic DNA was extracted from embryonic stem cells, bone marrow (BM) cells, and BMESL cells (clones 1 and 2). DNA was amplified using microsatellite

primers detecting polymorphisms between the bone marrow genome and the embryonic stem cell genome, separated on 5% agarose gel and visualized by ethidium bromide staining. Hybrid 129/B6 genotypes were detected in BMESL clones 1 and 2 for chromosomes 1 (D1MIT15), 9 (D9MIT48), 11 (D11MIT20), 14 (D14MIT11), 17 (D17MIT42) and 18 (D18MIT14).

phenotype on induced cell fusion of embryonic stem cells (or embryonic germ cells) and somatic cells has been demonstrated previously^{20,21}.

In four independent experiments, we reproducibly generated hyper-diploid embryonic stem-like cells, as shown above, when we treated a mixture of bone marrow and embryonic stem cells with either recombinant IL-3 or IL-3-conditioned medium (2–11 clones per experiment starting from 10^6 bone marrow cells). In contrast, we have not obtained any BMESL cells so far without addition of IL-3. Here, we used whole mononuclear bone marrow cells, and, at this moment, we do not know what fraction of cells in bone marrow is responsible for spontaneous cell fusion. The $\text{Sca1}^+ \text{Lin}^-$ fraction did not increase in the frequency of the hybrid cells (data not shown), suggesting that the haematopoietic stem cell itself is unlikely to be involved in cell fusion. It is known that spontaneous cell fusion occurs in macrophage cultures, forming multinucleated giant cells^{22–24}. Formation of multinucleated giant cells in cultures of monocytes and macrophages *in vitro* is enhanced by various cytokines including IL-3 (ref. 25). Taken together, macrophages and monocytes, or their precursors in bone marrow, might be involved in spontaneous cell fusion. Finally, the GFP mice (TgN(GFPU)5Nagy) were free of pathogens and healthy, indicating that they were not infected by lethal Sendai virus.

Our data demonstrate that mouse bone marrow cells can fuse spontaneously with other cells and subsequently adopt the phenotype of the recipient cells. This finding should be particularly significant when considering recent reports of transplanted bone marrow cells turning into unexpected cell types *in vivo*^{5–11}. Those studies discuss transdifferentiation on the basis of the existence of a Y chromosome (when the donor is male and the recipient is female) or the expression of other donor cell-specific genes. Our data could raise an alternative hypothesis in which donor bone marrow cells fuse with somatic cells *in vivo* as well and contribute to some of the apparently 'donor-typed' cells found in various organs. It should be noted that the frequency of spontaneous cell fusion was very low in the present study (2–11 clones per 10^6 bone marrow cells). Some *in*

vivo transplantation studies have reported robust (30–50%) levels of transdifferentiation⁸, which makes it unlikely that the results are due to cell fusion events. However, 'fused cells' also could become a dominant population when they had a growth or survival advantage over their parental cells by supplementing deficient genes. Contribution of cell fusion to apparently transdifferentiated cells *in vivo* is currently pure speculation; however, our data raise a warning to the overzealous trend in stem cell research to conclude transdifferentiation or dedifferentiation of cells without careful examination of genotypes. □

Methods

Bone marrow cell preparation

Mouse bone marrow cells were obtained from femurs of 8–9-week-old transgenic mice (female) expressing green fluorescence protein (GFP) (TgN(GFPU)5Nagy, Jackson Laboratories)¹². Bone marrow mononuclear cells were prepared by Ficoll–Hypaque gradient centrifugation.

Cell culture

The embryonic stem cell line R1 (129/Sv strain) was maintained on gelatin-coated tissue culture dishes in the following embryonic stem culture medium: DMEM (GIBCO) containing 15% FCS (Atlanta), 2 mM L-glutamine, 100 U ml^{−1} penicillin, 100 µg ml^{−1} streptomycin, 25 mM Hepes buffer (GIBCO), 300 µM monothioglycerol (Sigma), and 1,000 U ml^{−1} recombinant mouse LIF (ESGRO, CHEMICON), as we described previously¹³. Recombinant IL-3 was obtained from PeproTech. Puromycin was obtained from Sigma.

In vitro differentiation of BMESL cells

In vitro differentiation of BMESL cells was performed as we described before for embryonic stem cells^{14,15}. Briefly, BMESL cells were suspended in IMDM containing 2 mM L-glutamine, 100 U ml^{−1} penicillin, 100 µg ml^{−1} streptomycin, 20% FCS and 300 µM monothioglycerol (Sigma). Cells were cultured for 2 d by the hanging-drop method (1×10^3 embryonic stem cells per 30 µl in each drop). Embryoid bodies in hanging drops were transferred to suspension culture in Petri dishes and cultured for an additional 3 d. The resulting embryoid bodies were plated on to six-well tissue culture dishes for further differentiation.

RT-PCR

Total RNA was extracted by using an RNA aqueous kit (Ambion). Complementary DNA was synthesized from 2 µg total RNA, by using SuperScript II first-strand synthesis system with oligo(dT) (GIBCO). For each gene, the DNA primers were derived from different

exons to ensure that the PCR product represents the specific messenger RNA species and not genomic DNA. PCR was performed using Taq DNA polymerase (Eppendorf). The DNA sequence of the primers is summarized in Supplementary Information.

DNA ploidy and polymorphism

DNA contents per cell were determined by staining cells with propidium iodide and subsequent FACS analysis.

Genomic DNA were extracted from embryonic stem cells, bone marrow cells, and BMESL cells. DNA was amplified using microsatellite primers (D9MIT48, Research Genetics) detecting polymorphisms between the bone marrow genome (mostly C57BL/6) and the embryonic stem cell genome (129/Sv), separated on 5% agarose gel and visualized by ethidium bromide staining.

Received 17 January; accepted 27 February 2002.

Published online 13 March 2002, DOI 10.1038/nature730.

- Weissman, I. L. Translating stem and progenitor cell biology to the clinic: barriers and opportunities. *Science* **287**, 1442–1446 (2000).
- Wilmot, L., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810–813 (1997).
- Clarke, D. L. *et al.* Generalized potential of adult neural stem cells. *Science* **288**, 1660–1663 (2000).
- Pittenger, M. F. *et al.* Multilineage potential of adult human mesenchymal stem cells. *Science* **284**, 143–147 (1998).
- Ferrari, G. *et al.* Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* **279**, 1528–1530 (1998).
- Petersen, B. E. *et al.* Bone marrow as a potential source of hepatic oval cells. *Science* **284**, 1168–1170 (1999).
- Theise, N. D. *et al.* Liver from bone marrow in human. *Hepatology* **32**, 11–16 (2000).
- Lagasse, E. *et al.* Purified hematopoietic stem cells can differentiate into hepatocytes *in vivo*. *Nature Med.* **6**, 1229–1234 (2000).
- Brazelton, T. R., Rossi, F. M. V., Keshet, G. I. & Blau, H. M. From marrow to brain: expression of neuronal phenotypes in adult mice. *Science* **290**, 1775–1779 (2000).
- Mezey, E., Chandross, K. J., Harta, G., Maki, R. A. & McKecher, S. R. Turning blood into brain: cells bearing neuronal antigens generated *in vivo* from bone marrow. *Science* **290**, 1779–1782 (2000).
- Krause, D. S. *et al.* Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* **105**, 369–377 (2001).
- Hadjantonakis, A. K., Gertsenstein, M., Ikawa, M., Okabe, M. & Nagy, A. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech. Dev.* **76**, 79–90 (1998).
- Kawasome, H. *et al.* Targeted disruption of p70^{sk} defines its role in protein synthesis and rapamycin sensitivity. *Proc. Natl Acad. Sci. USA* **95**, 5033–5038 (1998).
- Minamino, T. *et al.* MEKK1 suppresses oxidative stress-induced apoptosis of embryonic stem cell-derived cardiac myocytes. *Proc. Natl Acad. Sci. USA* **96**, 15127–15132 (1999).
- Hamazaki, T. *et al.* Hepatic maturation in differentiating embryonic stem cells *in vitro*. *FEBS Lett.* **497**, 15–19 (2001).
- Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).
- Nishimoto, M., Fukushima, A., Okuda, A. & Muramatsu, M. The gene for the embryonic stem cell coactivator UTE1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell Biol.* **19**, 5453–5465 (1999).
- Kawasaki, H. *et al.* Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity. *Neuron* **28**, 31–40 (2000).
- Everett, C. A. & West, J. D. The influence of ploidy on the distribution of cells in chimeric mouse blastocysts. *Zygote* **4**, 59–66 (1996).
- Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr. Biol.* **11**, 1553–1558 (2001).
- Tada, M., Tada, T., Lefebvre, L., Barton, S. C. & Syran, M. A. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO J.* **16**, 6510–6520 (1997).
- Parwaresch, M. R., Kreipe, H. & Radzun, H. J. Human macrophage hybrid forming spontaneous giant cells. *Virchows Arch. B* **51**, 89–96 (1986).
- Chiozzi, P. *et al.* Spontaneous cell fusion in macrophage cultures expressing high levels of the P2Z/P2X7 receptor. *J. Cell Biol.* **138**, 697–706 (1997).
- Falzone, S. *et al.* The purinergic P2Z receptor of human macrophage cells. Characterization and possible physiological role. *J. Clin. Invest.* **95**, 1207–1216 (1995).
- Enelow, R. L., Sullivan, G. W., Carper, H. T. & Mandell, G. L. Induction of multinucleated giant cell formation from *in vitro* culture of human monocytes with interleukin-3 and interferon- β : comparison with other stimulating factors. *Am. J. Resp. Cell Mol. Biol.* **6**, 57–62 (1992).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

The authors are indebted to A. Nagy for providing embryonic stem cell lines; Y. Yoneda for discussion; J. Crawford, D. Steindler, S. May and S. Sugrue for critical reading of the manuscript; and G. Brown, M. Jorgenson, A. Meacham, N. Devine, and Diagnostic Cytogenetics for technical assistance. This work was supported by grants from the National Institutes of Health (to N.T. and E.W.S.).

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to N. T. (e-mail: terada@pathology.ufl.edu).

Changing potency by spontaneous fusion

Qi-Long Ying*, Jennifer Nichols*, Edward P. Evans† & Austin G. Smith*

* Centre for Genome Research, University of Edinburgh, The King's Buildings, West Mains Road, Edinburgh EH9 3JQ, UK

† Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Recent reports have suggested that mammalian stem cells residing in one tissue may have the capacity to produce differentiated cell types for other tissues and organs^{1–9}. Here we define a mechanism by which progenitor cells of the central nervous system can give rise to non-neural derivatives. Cells taken from mouse brain were co-cultured with pluripotent embryonic stem cells. Following selection for a transgenic marker carried only by the brain cells, undifferentiated stem cells are recovered in which the brain cell genome has undergone epigenetic reprogramming. However, these cells also carry a transgenic marker and chromosomes derived from the embryonic stem cells. Therefore the altered phenotype does not arise by direct conversion of brain to embryonic stem cell but rather through spontaneous generation of hybrid cells. The tetraploid hybrids exhibit full pluripotent character, including multilineage contribution to chimaeras. We propose that transdetermination consequent to cell fusion¹⁰ could underlie many observations otherwise attributed to an intrinsic plasticity of tissue stem cells⁹.

The capacity to generate differentiated cell types representing all three definitive germ layers has traditionally been considered a property reserved to cells of the inner cell mass and epiblast in the early embryo and derivative embryonal carcinoma and embryonic stem (ES) cells *in vitro*^{11,12}. Recently, however, remarkable potencies have been ascribed to stem cells isolated from various fetal and adult tissues including the central nervous system (CNS). CNS stem cells have been reported to contribute to haematopoietic lineages when injected into irradiated mice⁶, to produce muscle when co-cultured with skeletal myoblasts¹³, and to colonize multiple fetal lineages when introduced into pre-implantation embryos⁵. A further circumstance in which nervous system stem cells seem to change determination is during co-culture with differentiating ES cells, when they form multinucleated myotubes⁵. We investigated the basis of these phenomena and in particular whether during co-culture CNS cells may first convert to pluripotent ES cells as a route to generating other cell types.

We used distinct transgenic markers to isolate and identify descendants of both the ES cells and the CNS cells (Fig. 1). ZIN40 mice constitutively co-express resistance to the selection agent G418 and nuclear β -galactosidase activity¹⁴. Oct4-GiP mice express puromycin resistance and green fluorescent protein (GFP) exclusively in pluripotent and germline cells under direction of regulatory sequences of the mouse *Oct4* gene¹⁵. Neurosphere cultures⁵ were initiated from dissociated forebrains of transgenic fetuses at embryonic day 14.5 carrying either of these markers. Neurospheres were expanded for 3 days and then combined with HT2 ES cells. HT2 cells carry the hygromycin phosphotransferase–herpes simplex virus (HSV) thymidine kinase fusion gene (*Hytk*)¹⁶ inserted into the *Oct4* locus by homologous recombination^{17,18}. Selection was subsequently applied with G418 or puromycin as appropriate to eliminate the HT2 cells. After 2–4 weeks, proliferating cells with undifferentiated ES cell morphology re-emerged (Fig. 2b). Colonies were recovered from each of 23 independent co-cultures. These included one experiment in which fetal telencephalic cells were combined directly with ES cells without prior preparation of

neurospheres. Cells surviving selection expressed nuclear β -galactosidase (ZIN40 cells) or cytoplasmic GFP (Oct4-GiP cells) as appropriate. Expression of the Oct4-GiP transgene (Fig. 1b) is striking because, like Oct4 itself, this transgene is expressed only in germline and pluripotent cells and is not active in CNS cells *in vivo* or in primary culture (ref. 15 and our unpublished observations). GFP activity is therefore indicative of at least partial epigenetic reprogramming of the somatic cell genome to a state of pluripotency. Thus it appeared that ES cells could be generated from fetal brain cells.

We then tested these cells for expression of the hygromycin resistance marker carried only by HT2 ES cells (Fig. 1a). The cultures were completely resistant to hygromycin. Furthermore, they were sensitive to ganciclovir, which kills only cells harbouring the HSV *Tk* transgene. To exclude the possibility that inefficient selection could have given rise to mixed cultures, we generated several clonal derivatives by limiting dilution. The clonal isolates all exhibited fetal and ES cell resistance/sensitivity phenotypes. Therefore transgenic selection markers of different origin are co-expressed in individual cells.

We inferred that the selected cells were likely to be hybrids. Consistent with this, they had enlarged nuclei with multiple nucleoli (Fig. 2a, b). To test this interpretation directly, we prepared metaphase chromosome spreads (Fig. 2c). Analysis of 18 independent isolates revealed a tetraploid or near-tetraploid complement in all cases. The ES cells are male, and where a female fetus had been used as the source of brain cells, the sex chromosome complement was XXXY. The strain origin of chromosomes 8 and 14 can readily be discriminated owing to polymorphisms in centromeric heterochromatin¹⁹. Both strain 129 chromosomes of ES cell origin and non-strain 129 chromosomes of transgenic mouse origin were present in three separate isolates examined in this way (Fig. 2d).

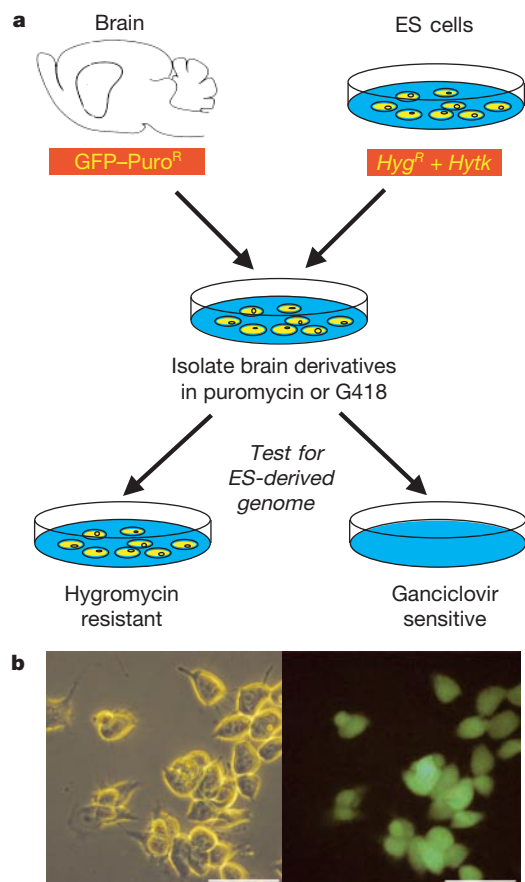


Figure 1 Isolation and identification of descendants of ES and CNS cells. **a**, Diagram of experimental strategy. **b**, Expression of GFP in undifferentiated cells isolated after co-culture of Oct4-GiP fetal brain cells with HT2 ES cells and selection in puromycin.

These observations can be explained only by the formation of cell hybrids between CNS and ES cells.

The hybrid cells expressed the ES cell markers¹² alkaline phosphatase and stage-specific embryonic antigen (SSEA)-1 (data not shown). They expressed the essential pluripotent cell transcription factor Oct-4^{18,20} as well as the Oct4-GiP transgene. Like normal ES cells, they were dependent on the self-renewal factor leukaemia inhibitory factor (LIF) to suppress differentiation¹². On aggregation they formed embryoid bodies²¹ containing both extra-embryonic endoderm and spontaneously contracting cardiomyocytes (Fig. 3). Embryoid bodies treated with retinoic acid²² gave rise to neurons. Thus the hybrids have *in vitro* self-renewal and differentiation properties similar to those of regular ES cells¹². Hybrids produced by electrofusion of ES cells with thymocytes have similarly been found to exhibit ES cell characteristics²³.

We examined the capacity for incorporation into embryonic development by blastocyst injection. Contributions to fetal tissues were detected in 8 of 23 transferred embryos by β -galactosidase staining for the ZIN40 marker (Fig. 4a). The contributions were modest by comparison with standard ES cells and uneven between tissues, but this is expected owing to competitive overgrowth of tetraploid hybrid cells by diploid host cells²⁴. Interestingly, 1 of 14 live-born mice showed overt coat-colour chimaerism (Fig. 4b). This animal was euthanized and we analysed the internal organs for expression of nuclear β -galactosidase. Staining was detected in intestine, kidney, heart and most prominently in liver (Fig. 4c). A proportion of tetraploid hepatocytes is normally present in the liver, which may account for the persistent contribution of hybrid cells in this organ. These observations establish that the presence of a genome derived from a brain cell is not prohibitive for multilineage contribution of ES cell hybrids.

Finally we investigated whether similar fusion events could occur with cells isolated from adult brain. A modified schedule for co-culture and selection was adopted to allow for the comparatively slower proliferation rate of adult-derived neurospheres. Neurospheres were established from transgenic mice that express ubiqui-

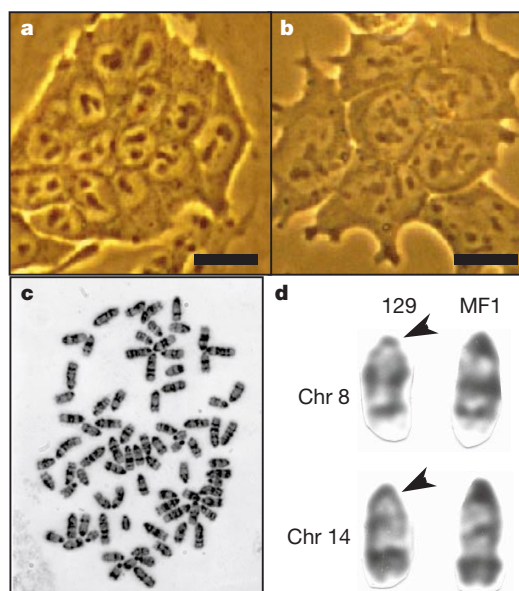


Figure 2 Morphology and chromosomal constitution. **a**, **b**, Photomicrographs of parental ES cells (**a**) and cells isolated after co-culture (**b**) showing enlarged nuclei with multiple nucleoli in cells with otherwise typical ES cell morphology. Scale bar, 25 μ m. **c**, Metaphase spread showing tetraploid chromosome complement after co-culture of adult brain neurospheres and ES cells. **d**, Strain 129 and non-strain 129 chromosomes in cells derived from co-culture. Arrowheads indicate the polymorphic centromeric C bands that are diagnostically reduced in strain 129 chromosomes and large in MF1 or C57BL/6 chromosomes of the transgenic mice.

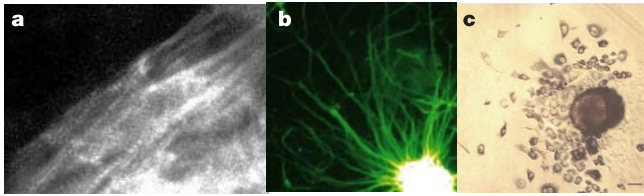


Figure 3 Pluripotency of hybrid cells. Embryoid bodies were prepared from hybrid cells and outgrowths analysed by immunostaining for markers of cardiomyocytes (β -actin, **a**), or, if treated with retinoic acid, neurons (type III tubulin, **b**), or by *in situ* hybridization (**c**), which revealed *Sparc* messenger RNA in migratory cells of typical parietal endoderm morphology.

tously the tauGFP fusion protein linked with puromycin resistance²⁵. The lateral ventricles of an 8-week-old female mouse were dissociated and neurospheres were propagated for 2 weeks as described⁵. These neurosphere cells were combined in a 10:1 ratio with ES cells in adherent culture. The co-culture was maintained in neural stem cell medium supplemented with LIF for 1 week. The culture was then transferred to ES cell medium and expanded for a further week before applying puromycin selection. After 2 weeks of selection, proliferating colonies of tauGFP-expressing ES cells appeared. These cells had large nuclei and multiple nucleoli as above. Chromosome preparations confirmed that they were tetraploid with an XXXY complement (Fig. 2c) and contained both strain 129 and non-strain 129 chromosomes 8 and 14.

The frequency of hybrid isolation in these experiments is between 10^{-4} and 10^{-5} per brain cell plated, although note that we have made no efforts at optimization. There is no fusogen present under our culture conditions, implying that the hybrids are generated by spontaneous fusion between ES cells and CNS cells. The phenomenon is not dependent on the use of HT2 cells because an independently derived ES cell line, 46C, was used in the fusion with adult neurosphere cells. We have also obtained hybrids from co-cultures between ES cells. Spontaneous fusion between mammalian cells was originally documented in 1961^{26,27}. This provided the foundation for development of somatic cell genetics and the appreciation that cell specification can be reprogrammed in

hybrids¹⁰. This work has been overlooked in recent reports of transdetermination and transdifferentiation^{1–8,13}, resulting in a major challenge to the concept of progressive lineage restriction during development⁹. However, as we have shown here, the fact that mammalian cells can spontaneously form hybrids presents an alternative explanation for the apparent ability of cells from one tissue to generate progeny of another type. Future claims of cell plasticity should therefore examine the presence not just of donor cell markers but also of host cell markers in any putative trans-determination product. □

Methods

Embryonic stem cells and hybrid cells were maintained in LIF-supplemented medium without feeders²⁸. HT2 and 46C ES cells are of pure inbred 1290la origin. Transgenic mice are on mixed 129 $\times$ MF1 (ZIN40 and Oct4-GiP) and 129 $\times$ C57BL/6 (tauGFP) backgrounds. Neurosphere cultures were initiated and expanded in DMEM/F12 medium with N2 and B27 supplements plus fibroblast growth factor-2 (10 ng ml^{-1})⁵. Dissociated fetal neurosphere cells and ES cells were mixed in a 2:1 ratio and plated on dishes coated with poly-D-lysine and laminin at $1\text{--}1.5 \times 10^4 \text{ cells cm}^{-2}$. Cultures were maintained in neural stem cell medium for 48 h, then changed to ES cell medium containing fetal calf serum and LIF. G418 ($200\text{--}400 \mu\text{g ml}^{-1}$) was added after a further 2 days of co-culture, but puromycin ($1 \mu\text{g ml}^{-1}$) only after 6 days because the Oct4-GiP transgene is not expressed in brain cells. In the case of adult neurosphere cells, the co-culture was initiated at a 10:1 ratio of brain cells to ES cells and maintained for 2 weeks before application of puromycin selection. ES cell differentiation and chimera production followed standard protocols²⁹.

Received 28 November 2001; accepted 27 February 2002.

Published online 13 March 2002; DOI 10.1038/nature729.

1. Brazelton, T. R., Rossi, F. M., Keshet, G. I. & Blau, H. M. From marrow to brain: expression of neuronal phenotypes in adult mice. *Science* **290**, 1775–1779 (2000).
2. Mezey, E., Chandross, K. J., Harta, G., Maki, R. A. & McKecher, S. R. Turning blood into brain: cells bearing neuronal antigens generated *in vivo* from bone marrow. *Science* **290**, 1779–1782 (2000).
3. Krause, D. S. *et al.* Multi-organ, multi-lineage engraftment by a single bone marrow-derived stem cell. *Cell* **105**, 369–377 (2001).
4. Lagasse, E. *et al.* Purified hematopoietic stem cells can differentiate into hepatocytes *in vivo*. *Nature Med.* **6**, 1229–1234 (2000).
5. Clarke, D. L. *et al.* Generalized potential of adult neural stem cells. *Science* **288**, 1660–1663 (2000).
6. Bjornson, C. R., Rietze, R. L., Reynolds, B. A., Magli, M. C. & Vescovi, A. L. Turning brain into blood: a hematopoietic fate adopted by adult neural stem cells *in vivo*. *Science* **283**, 534–537 (1999).
7. Ferrari, G. *et al.* Muscle regeneration by bone marrow-derived myogenic progenitors. *Science* **279**, 1528–1530 (1998).
8. Orlie, D. *et al.* Bone marrow cells regenerate infarcted myocardium. *Nature* **410**, 701–705 (2001).
9. Blau, H. M., Brazelton, T. R. & Weimann, J. M. The evolving concept of a stem cell: entity or function? *Cell* **105**, 829–841 (2001).
10. Ephrussi, B. *Hybridization of Somatic Cells* (Princeton Univ. Press, Princeton, 1972).
11. Gardner, R. L. & Beddington, R. S. Multi-lineage 'stem' cells in the mammalian embryo. *J. Cell Sci.* **10** (Suppl.), 11–27 (1988).
12. Smith, A. in *Stem Cell Biology* (ed. Marshak, D. R., Gardner, R. L., Gottlieb, D.) 205–230 (Cold Spring Harbor Laboratory Press, New York, 2001).
13. Galli, R. *et al.* Skeletal myogenic potential of human and mouse neural stem cells. *Nature Neurosci.* **3**, 986–991 (2000).
14. Mountford, P. & Smith, A. G. Internal ribosome entry sites and dicistronic RNAs in mammalian transgenesis. *Trends Genet.* **11**, 179–184 (1995).
15. Yeom, Y. I. *et al.* Germ-line regulatory element of Oct-4 specific for the totipotent cycle of embryonic cells. *Development* **122**, 881–894 (1996).
16. Lupton, S. D., Brunton, L. L., Kalberg, V. A. & Overell, R. W. Dominant positive and negative selection using a hygromycin phosphotransferase–thymidine kinase fusion gene. *Mol. Cell. Biol.* **11**, 3374–3378 (1991).
17. Mountford, P. *et al.* Dicistronic targeting constructs: reporters and modifiers of mammalian gene expression. *Proc. Natl Acad. Sci. USA* **91**, 4303–4307 (1994).
18. Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet.* **24**, 372–376 (2000).
19. Akesson, E. C. & Davidson, M. T. in *Genetic Variants and Strains of the Laboratory Mouse* (eds Lyon, M., Rastan, S. & Brown, S.) 1506–1509 (Oxford Univ. Press, Oxford, 1996).
20. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct-4. *Cell* **95**, 379–391 (1998).
21. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morphol.* **87**, 27–45 (1985).
22. Bain, G., Kitchens, D., Yao, M., Huettner, J. E. & Gottlieb, D. I. Embryonic stem cells express neuronal properties *in vitro*. *Dev. Biol.* **168**, 342–357 (1995).
23. Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. & Tada, T. Nuclear reprogramming of somatic cells by *in vitro* hybridization with ES cells. *Curr. Biol.* **11**, 1553–1558 (2001).
24. Nagy, A. *et al.* Embryonic stem cells alone are able to support fetal development in the mouse. *Development* **110**, 815–821 (1991).
25. Pratt, T., Sharp, L., Nichols, J., Price, D. J. & Mason, J. O. Embryonic stem cells and transgenic mice ubiquitously expressing a tau-tagged green fluorescent protein. *Dev. Biol.* **228**, 19–28 (2000).
26. Barski, G., Sorieul, S. & Cornet, F. "Hybrid" type cells in combined cultures of two different mammalian cell strains. *J. Natl Cancer Inst.* **26**, 1269–1291 (1961).
27. Sorieul, S. & Ephrussi, B. Karyological demonstration of hybridization of mammalian cells *in vitro*. *Nature* **190**, 653–654 (1961).

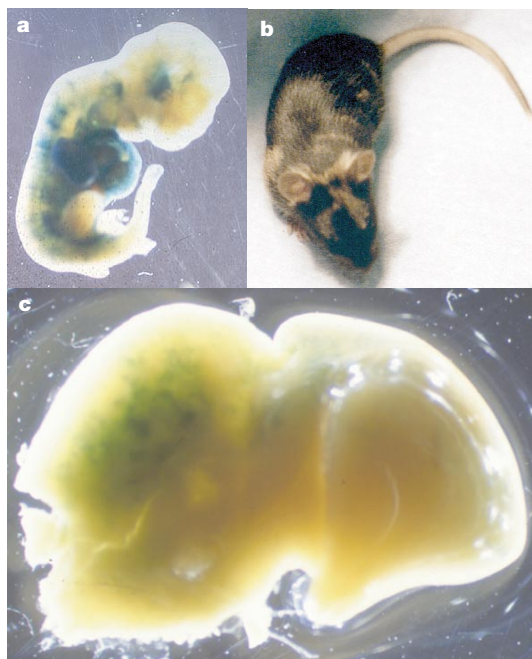


Figure 4 Contribution of ZIN40/HT2 hybrid cells to chimaeras. **a**, β -galactosidase staining of fetal chimaera at embryonic day 11.5, slightly retarded. **b**, Adult mouse chimaera. **c**, β -galactosidase staining of liver from an adult mouse chimaera.

28. Smith, A. G. Culture and differentiation of embryonic stem cells. *J. Tissue Culture Methods* **13**, 89–94 (1991).
29. Robertson, E. J. *Teratocarcinoma and Embryo-derived Stem Cells: A Practical Approach* (IRL, Oxford, 1987).

Acknowledgements

H. Niwa and I. Chambers generated the HT2 ES cells and Marios Stavridis the 46C cells. We thank C. Graham for directing us to the original studies of Barski and Ephrussi and are grateful to C. Blackburn and A. Medvinsky for comments on the manuscript. This research was supported by the International Human Frontiers Science Program and by the Medical Research Council and Biotechnology and Biological Sciences Research Council of the UK.

Competing interests statement

The authors declare competing financial interests: see the website (<http://www.nature.com>) for details.

Correspondence and requests for materials should be addressed to A.G.S.
(e-mail: austin.smith@ed.ac.uk).

Shaggy/GSK3 antagonizes Hedgehog signalling by regulating Cubitus interruptus

Jianhang Jia^{*†}, Kazuhito Amanai^{*†}, Gelin Wang^{*†}, Jiong Tang[‡],
Bing Wang^{*} & Jin Jiang^{*}

^{*} Center for Developmental Biology and Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9133, USA
[‡] Division of Cell and Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9133, USA
[†] These authors contributed equally to this work

The *Drosophila* protein Shaggy (Sgg, also known as Zeste-white3, Zw3) and its vertebrate orthologue glycogen synthase kinase 3 (GSK3) are inhibitory components of the Wingless (Wg) and Wnt pathways¹. Here we show that Sgg is also a negative regulator in the Hedgehog (Hh) pathway. In *Drosophila*, Hh acts both by blocking the proteolytic processing of full-length Cubitus interruptus, Ci (Ci155), to generate a truncated repressor form (Ci75), and by stimulating the activity of accumulated Ci155 (refs 2–6). Loss of *sgg* gene function results in a cell-autonomous accumulation of high levels of Ci155 and the ectopic expression of Hh-responsive genes including *decapentaplegic* (*dpp*) and *wg*. Simultaneous removal of *sgg* and *Suppressor of fused*, *Su(fu)*⁷, results in wing duplications similar to those caused by ectopic Hh signalling. Ci is phosphorylated by GSK3 after a primed phosphorylation by protein kinase A (PKA), and mutating GSK3-phosphorylation sites in Ci blocks its processing and prevents the production of the repressor form. We propose that Sgg/GSK3 acts in conjunction with PKA to cause hyperphosphorylation of Ci, which targets it for proteolytic processing, and that Hh opposes Ci proteolysis by promoting its dephosphorylation.

The Hh family of secreted proteins controls cell growth and patterning in many principal developmental processes in both vertebrates and invertebrates⁸. Moreover, mutations in components of the Hh signalling pathway have been implicated in many human disorders, including cancer⁹. During *Drosophila* limb development, posterior (P)-compartment cells express and secrete Hh that induces adjacent anterior (A)-compartment cells to express target genes including *dpp*, *wg* (leg only) and *patched* (*ptc*) by regulating the transcription factor Ci^{4,10,11}. In A-compartment cells distant from the AP compartment boundary, Ci is processed to generate a truncated repressor form (Ci75) that represses a subset of Hh-

responsive genes including *dpp*^{2,4}. In A-compartment cells adjacent to the AP compartment border, Hh signalling blocks Ci processing to generate Ci75, and causes the accumulation of full-length Ci (Ci155)². In addition, high levels of Hh stimulate a distinct transcriptional activation activity of Ci155, which is required for the expression of Hh-responsive genes such as *ptc*^{3,4,12}.

In both wing and leg discs, loss of *sgg* function in the A compartment either by using *sgg* mutations or by overexpressing a dominant negative form of GSK3 (DN-GSK3)¹³ causes the accumulation of high levels of Ci155 in a cell-autonomous fashion without affecting *ci-lacZ* expression (Fig. 1a, b, e, h, j–l, and see Supplementary Information). In wing discs, anterior *sgg*[−] cells or DN-GSK3-expressing cells located outside the wing pouch region ectopically express *dpp*, which is repressed by Ci75 (Fig. 1c, d). However, anterior *sgg*[−] cells do not ectopically activate *ptc*, which is activated by Ci155 (Fig. 1f). In leg discs, anterodorsal *sgg*[−] cells distant from the AP boundary ectopically express *wg* and low levels of *dpp* (Fig. 1g–i), a phenotype similar to that associated with *sgg* PKA double-mutant cells in which both Wg and Hh signalling pathways are ectopically activated¹⁴. As in the case of wing discs,

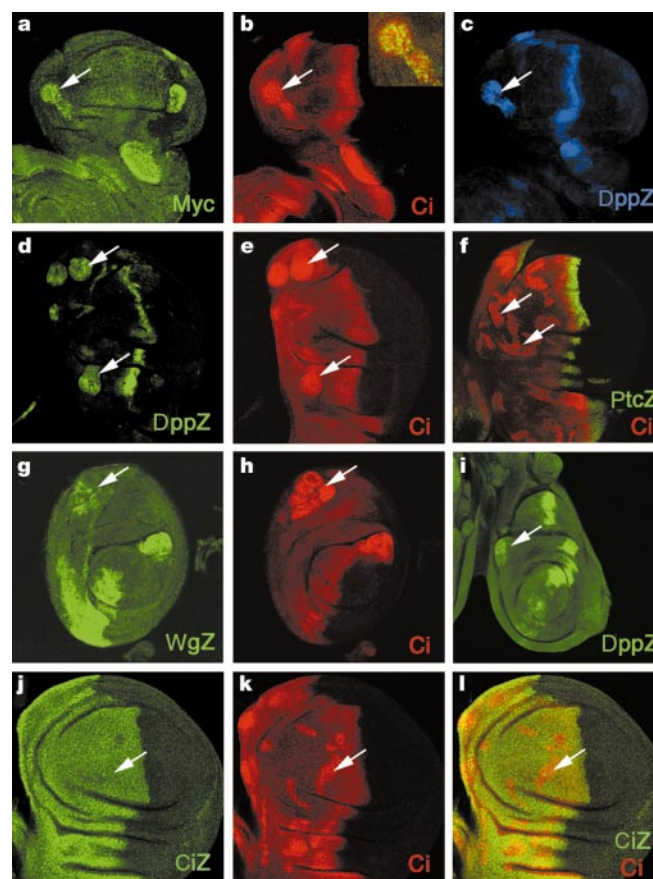


Figure 1 Ectopic Hh signalling activity in *sgg*[−] cells. **a–c**, A wing disc carrying *sgg*[−] clones and showing the expression of a marker gene (**a**), Ci155 (**b**) and *dpp-lacZ* (**c**). *sgg*[−] cells are recognized by strong Myc staining. Anterior *sgg*[−] cells accumulate high levels of Ci155 and ectopically express *dpp-lacZ*. The inset in **b** shows an enlarged view of the anterior *sgg*[−] clone (arrow in **a**). **d, e**, A wing disc containing clones of cells that express UAS-DN-GSK3 using *actin5c > CD2 > Gal4*. Anterior DN-GSK3-expressing cells accumulate high levels of Ci155 (arrows in **e**) and ectopically express *dpp-lacZ* (arrows in **d**). **f**, A wing disc carrying *sgg*[−] mutant clones and exhibiting the expression of Ci155 and *ptc-lacZ*. Anterior *sgg*[−] mutant clones do not ectopically activate *ptc-lacZ* (arrows). **g–i**, Leg discs carrying *sgg*[−] clones and showing the expression of *wg-lacZ* (**g**), Ci155 (**h**) and *dpp-lacZ* (**i**). Anterodorsal *sgg*[−] cells accumulate high levels of Ci155 (**h**) and ectopically express low levels of *wg-lacZ* (**g**) and *dpp-lacZ* (**i**). **j–l**, A wing disc carrying *sgg*[−] clones and showing the expression of *ci-lacZ* and Ci155. *sgg*[−] cells (arrow) do not express *ci-lacZ* at higher levels.

sgg⁻ cells seem to transduce low levels of Hh signalling activity, because *wg* is not fully activated (Fig. 1g), and little, if any, *ptc* is expressed (data not shown). One hypothesis that accounts for these observations is that loss of *sgg* function affects Ci processing to generate Ci75 but does not stimulate the activity of Ci155.

The activity of Ci155 is regulated by several mechanisms including attenuation by *Su(fu)*^{3,7}. Whereas loss of *Su(fu)* function does not cause any significant phenotypes, it dramatically enhances *sgg*⁻ phenotypes (Fig. 2). For example, anterior *sgg Su(fu)* double-mutant clones organize wing duplication (Fig. 2c, d), whereas *sgg* single-mutant clones form only small outgrowths (Fig. 2a, b). In wing discs, anterior *sgg Su(fu)* double-mutant cells activate low levels of *ptc* (Fig. 2e–g), which is not ectopically expressed in *sgg* single-mutant cells (Fig. 1f). In leg discs, anterior *sgg Su(fu)* double-mutant cells express *ptc* and high levels of *wg* (Fig. 2h–m). Hence,

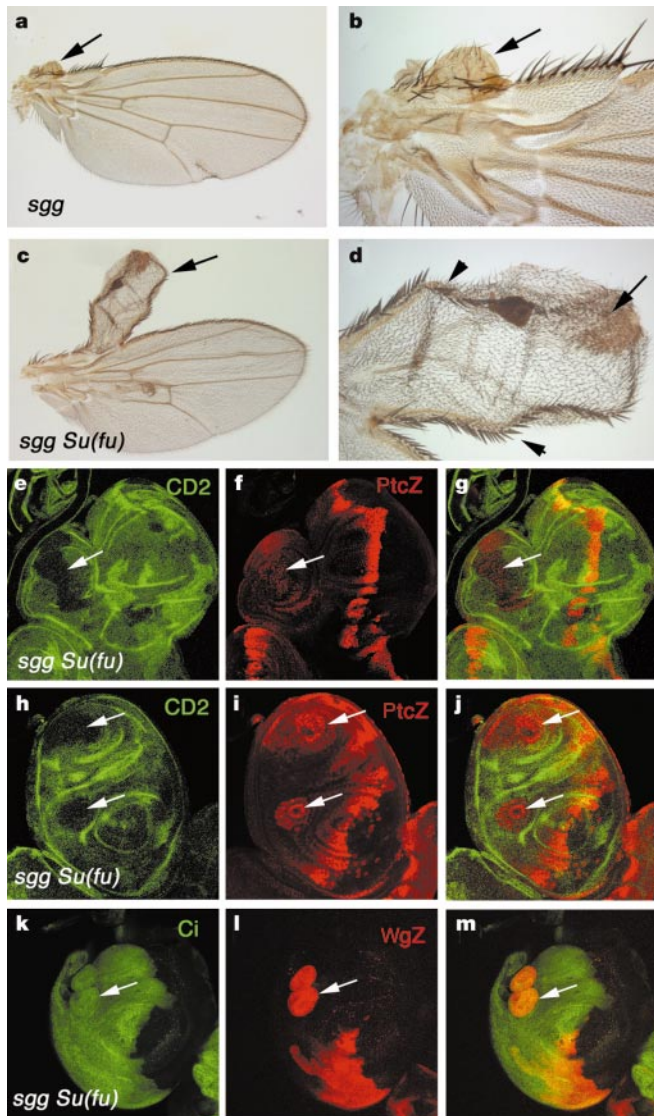


Figure 2 *Su(fu)* mutation enhances *sgg*⁻ phenotypes. **a, b**, A mosaic wing with an anterior *sgg*⁻ clone (marked by the yellow (*y*) mutation) that forms a small outgrowth. **c, d**, An anterior *sgg Su(fu)* double-mutant clone (marked by *y*) organizes the formation of a duplicated wing. *sgg Su(fu)* double-mutant cells form ectopic sensory bristles (arrow in **d**). Most of the duplicated wing is derived from wild-type cells marked by *y*⁺ (arrowheads in **d**). **e–j**, Mosaic wing (**e–g**) and leg (**h–j**) discs showing the expression of a marker gene (*hsp-CD2*) and *ptc-lacZ*. *sgg Su(fu)* double-mutant cells are marked by the absence of CD2 expression (arrows). Anterior *sgg Su(fu)* double-mutant cells activate *ptc-lacZ*, although the level of ectopic *ptc-lacZ* expression is low in the wing disc. **k–m**, In a mosaic leg disc, anterior *sgg Su(fu)* double-mutant cells activate *wg-lacZ* at high levels.

Ci155 accumulated in *sgg*⁻ cells is largely inactive and its activity is stimulated by removal of *Su(fu)* function.

Ectopic activation of the Wg pathway by overexpression of a constitutively active form of Armadillo (Δ Arm)¹⁵ did not cause the accumulation of high levels of Ci155 or activate any Hh-responsive genes (Fig. 3a–c, data not shown). Hence, the constitutive Hh signalling activity in *sgg*⁻ cells is not secondary to aberrant activation of the Wg pathway in these cells. Hh induces stabilization of Smoothened (Smo), a seven-transmembrane protein that transduces the Hh signal¹⁶, but anterior *sgg*⁻ cells do not stabilize Smo (Fig. 3d, e). In addition, *sgg smo* double-mutant cells, like *sgg* single-mutant cells, accumulate high levels of Ci155 (Fig. 3f), suggesting that Sgg acts downstream of Smo to regulate Ci processing. The proteolytic processing of Ci requires the activities of several intracellular Hh signalling components including PKA and the kinesin-related protein Costal2 (Cos2)^{6,12,17–20}. Overexpressing either Cos2 or a constitutively active form of PKA (mC*) blocks the accumulation of Ci155 induced by Hh^{6,21} (Fig. 3h–i). In contrast, wing discs overexpressing mC* or Cos2 accumulate high levels of Ci155 after treatment with 50 mM LiCl, a specific inhibitor of GSK3 kinase activity²² (Fig. 3k–l). These observations suggest that Sgg acts downstream of, or in parallel with, PKA and Cos2 to regulate Ci processing.

PKA promotes Ci processing by directly phosphorylating it at multiple sites in its carboxy-terminal region^{6,17,19}. We therefore

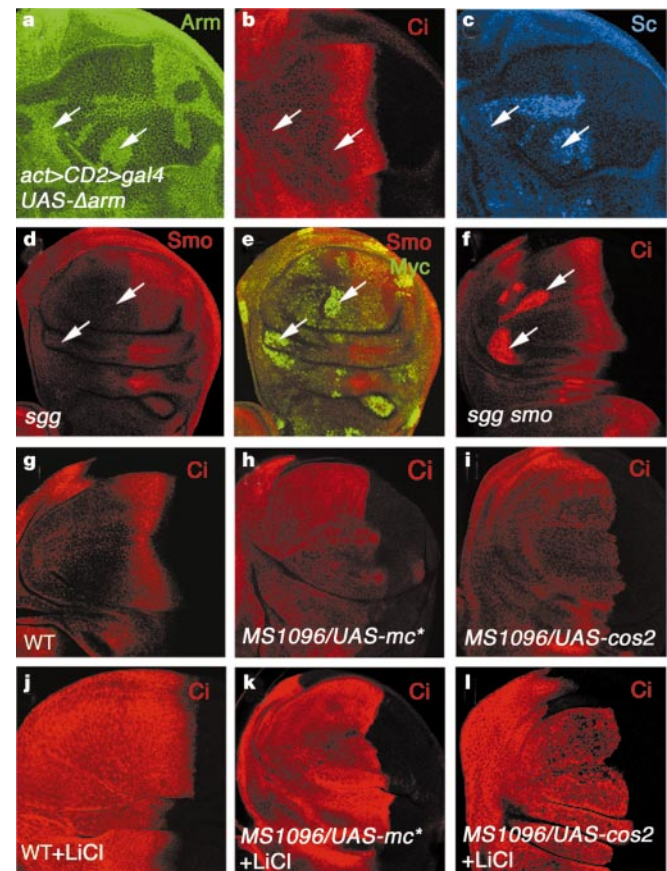


Figure 3 Sgg acts in the Hh pathway downstream of, or in parallel with, PKA and Cos2. **a–c**, Wing disc cells ectopically expressing Δ Arm (arrows) activate *scute* (*sc*), but do not cause the accumulation of Ci155. **d, e**, Anterior *sgg*⁻ cells (arrows) do not stabilize Smo protein. **f**, *sgg smo* double-mutant cells accumulate high levels of Ci155 (arrows). **g, j**, Wild-type (WT) wing discs treated with (j) or without (g) LiCl. LiCl treatment causes the accumulation of high levels of Ci155 throughout the A-compartment (j). **h, i, k, l**, Wing discs expressing *UAS-mC** (**h, k**) or *UAS-cos2* (**i, l**) and treated with (k, l) or without (h, i) LiCl. Overexpressing mC* or Cos2 blocks the accumulation of Ci155 induced by Hh (**h, i**), but does not block the accumulation of Ci155 caused by LiCl treatment (**k, l**).

investigated whether Sgg/GSK3 also regulates Ci processing by direct phosphorylation. The canonical GSK3-phosphorylation site consists of two Ser/Thr residues separated by three amino acids: (Ser/Thr0)-X-X-X-(Ser/Thr+4). Phosphorylation of Ser/Thr at the +4 position by a priming kinase allows GSK3 to phosphorylate Ser/Thr at the 0 position²³. Examination of Ci sequence revealed three GSK3 consensus sites (Ser 852, Ser 884 and Ser 888) adjacent to two previously identified PKA phosphorylation sites (Ser 856 and Ser 892)^{6,17,19} (Fig. 4a). To determine whether Ci can be a direct substrate of Sgg/GSK3, we carried out an *in vitro* kinase assay. Three glutathione *S*-transferase (GST)-Ci fusion proteins containing Ci fragments from amino acids 441–1,065 were generated: GST-Ci contains wild-type sequence; GST-Ci-3P has three PKA sites mutated (S838A, S856A and S892A); and GST-Cim3 has three GSK3 consensus sites mutated (S852A, S884A and S888A) (Fig. 4a). GST-Ci is specifically phosphorylated by PKA but not by GSK3 without prior PKA treatment (Fig. 4b, lanes 1, 2 and 4); however, it becomes a good substrate for GSK3 after primed phosphorylation by PKA (Fig. 4b, lane 3). In contrast, neither GST-Ci-3P nor GST-Cim3 can be phosphorylated by GSK3 even after PKA treatment (Fig. 4b, lanes 7 and 9). These results suggest that primed phosphorylation by PKA at Ser 856 and Ser 892 allows GSK3 to phosphorylate Ci at Ser 852, Ser 884 and Ser 888.

To determine whether Sgg/GSK3 phosphorylates Ci *in vivo*, we examined Ci phosphorylation in cl-8 cells treated with or without LiCl, which specifically blocks GSK3 kinase activity²². As shown in Fig. 4c and consistent with a previous report⁵, treating cl-8 cells with

both a proteasome inhibitor, MG132, and a phosphatase inhibitor, okadaic acid (OA), results in the accumulation of hyperphosphorylated forms of Ci155, which exhibit much slower electrophoretic mobility on SDS polyacrylamide gel than unphosphorylated or hypophosphorylated forms (Fig. 4c, lane 3). Treating cells with MG132 and OA in the presence of 50 mM LiCl results in hypophosphorylation of Ci because it eliminates the slowest-migrating forms of Ci155 (Fig. 4c, lane 4). These observations suggest that inhibition of Sgg/GSK3 kinase activity affects Ci phosphorylation in intact cells. The residual phosphorylation of Ci155 in the presence of LiCl is probably due to phosphorylation by PKA because LiCl does not inhibit PKA kinase activity²².

If Ci processing is regulated by GSK3 phosphorylation, one would predict that mutating GSK3-phosphorylation sites in Ci should block its processing to generate the Ci75 repressor. To test

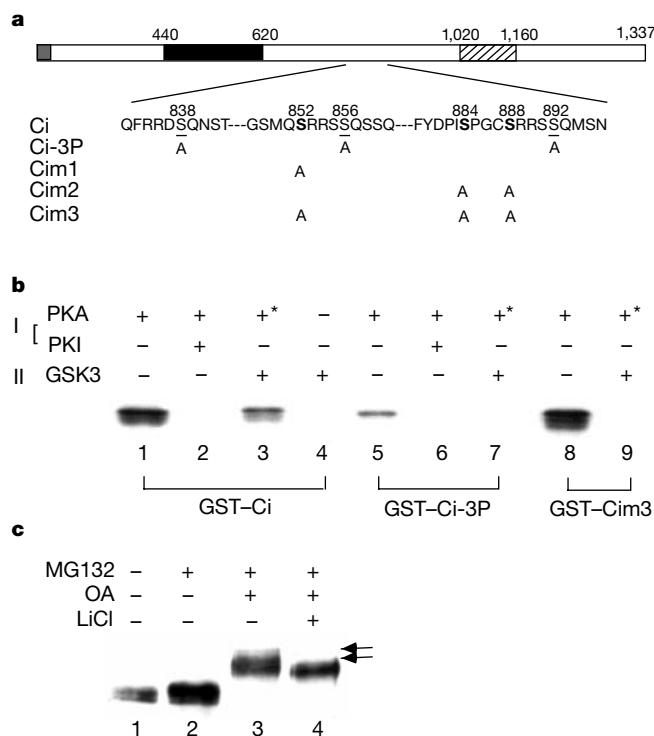


Figure 4 Phosphorylation of Ci by Sgg/GSK3. **a**, A schematic drawing of HA-Ci. Grey box, two copies of HA tag; filled box, the Zn-finger DNA-binding domain; hatched box, the *Drosophila* CBP-binding domain. The underlined residues are PKA-phosphorylation sites whereas the bold residues are the GSK3-phosphorylation consensus sites.

b, Autoradiograph of the *in vitro* kinase assay. GST fusion proteins were incubated with the PKA catalytic subunit (with or without PKI) or GSK3 in the presence of γ -³²P-ATP, or treated with GSK3 in the presence of γ -³²P-ATP and PKI after incubation with PKA in the presence of unlabelled ATP (indicated by an asterisk). **c**, cl-8 cells were treated with 50 μ M MG132 (lane 2) or a combination of 50 μ M MG132 and 10 nM OA in the absence (lane 3) or presence (lane 4) of 50 mM LiCl for 4 h before collection. Cell lysates were separated by SDS-polyacrylamide gel electrophoresis, followed by western blot analysis using the anti-Ci antibody 2A1. LiCl treatment eliminates the slowest-migrating forms of Ci (arrows).

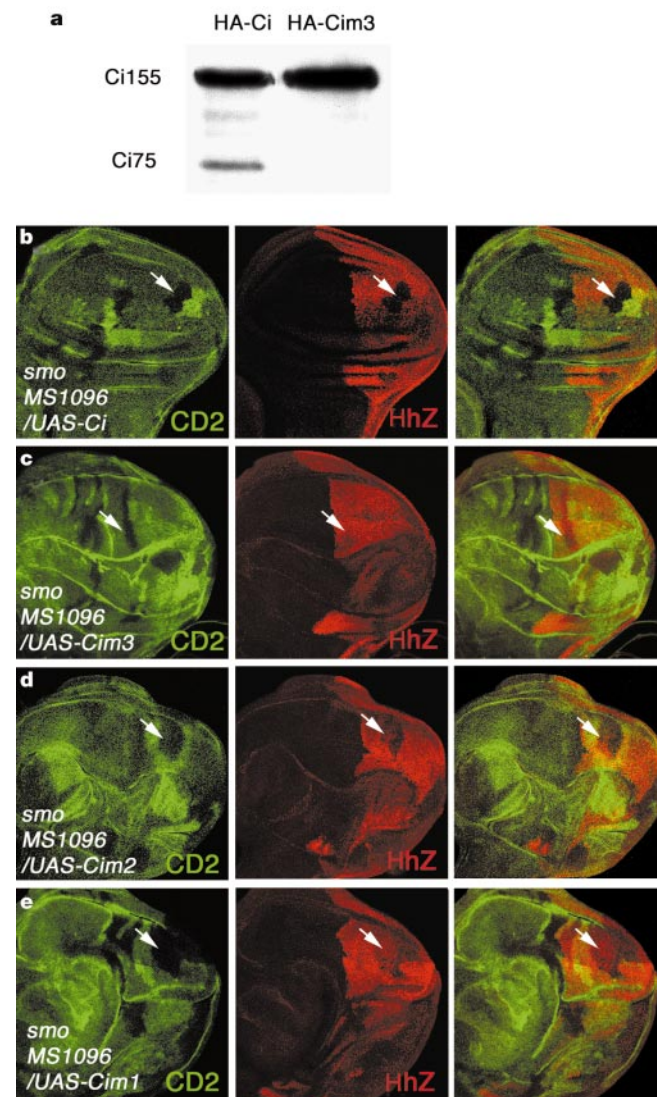


Figure 5 Mutating GSK3-phosphorylation sites in Ci impedes its processing to generate the repressor form. **a**, Western blot analysis of protein extracts from wing discs expressing HA-Ci or HA-Cim3 using a *MS1096* Gal4 driver. Protein extracts were prepared from about 400 wing discs, immunoprecipitated and blotted with an anti-HA antibody.

b-e, Wing discs expressing wild-type Ci (**b**), Cim3 (**c**), Cim2 (**d**) and Cim1 (**e**) under the control of the *MS1096* Gal4 driver. These discs also express *hh-lacZ* and contain *smo*⁻ cells (marked by the absence of CD2 staining). P-compartment *smo*⁻ cells expressing wild-type Ci block *hh-lacZ* expression (arrow in **b**), whereas *smo*⁻ cells expressing Cim3 do not (arrow in **c**). *smo*⁻ cells expressing Cim2 inhibit *hh-lacZ* expression to a lesser extent than those expressing Ci (arrow in **d**). *smo*⁻ cells expressing Cim1 only slightly attenuate *hh-lacZ* expression (arrow in **e**).

this, we generated UAS transgenes containing haemagglutinin (HA)-tagged wild-type (UAS-HA-Ci) or mutant Ci (UAS-HA-Cim3); these were expressed in wing discs using the Gal4/UAS system²⁴. Whereas HA-Ci is partially processed into Ci75, HA-Cim3 does not give rise to detectable Ci75 (Fig. 5a). We also examined the effect on the production of Ci75 of mutating GSK3-phosphorylation sites using an *in vivo* function assay⁴. We ectopically expressed either wild-type or mutant Ci in the P-compartment of wing discs carrying *smo*⁻ clones, and examined *hh-lacZ* expression, which is inhibited by Ci75 (ref. 4). P-compartment *smo*⁻ cells expressing HA-Ci block *hh-lacZ* expression (Fig. 5b), indicating that wild-type Ci is processed to generate Ci75 in the absence of Hh signalling. In contrast, P-compartment *smo*⁻ cells expressing HA-Cim3 do not repress *hh-lacZ* expression (Fig. 5c), indicating that HA-Cim3 does not produce Ci75 *in vivo*. Hence, mutating GSK3-phosphorylation sites in Ci affects its proteolytic processing to generate the repressor form.

To assess the importance of individual GSK3-phosphorylation sites for Ci processing, we examined two additional mutant forms of Ci—HA-Cim1 and HA-Cim2 (Fig. 4a)—using the *in vivo* function assay⁴. P-compartment *smo*⁻ cells expressing HA-Cim2 partially block *hh-lacZ* expression (Fig. 5d), suggesting that lack of phosphorylation at Ser884 and Ser888 attenuates Ci processing. P-compartment *smo*⁻ cells expressing HA-Cim1 also partially inhibit *hh-lacZ* expression, however, to a lesser extent than those expressing HA-Cim2 (Fig. 5e), suggesting that mutating Ser852 greatly impedes, although does not completely abolish, Ci processing. Hence, efficient processing of Ci seems to require phosphorylation at all three GSK3 sites.

Taken together, our data suggest that Sgg/GSK3 acts in conjunction with PKA to promote hyperphosphorylation of Ci, which is essential for efficient Ci processing to generate its repressor form. The requirement for multiple phosphorylation seems to be a general mechanism to regulate proteolysis of regulatory proteins²⁵. The involvement of multiple phosphorylation may provide a way for Hh to differentially regulate Ci by controlling its level of phosphorylation. For example, low levels of Hh may cause partial dephosphorylation of Ci by opposing Sgg, resulting in inhibition of Ci processing to generate Ci75 but leaving Ci155 inactive. In contrast, high levels of Hh may cause complete dephosphorylation of Ci by opposing PKA, which not only blocks Ci processing but also stimulates the activity of accumulated Ci155.

GSK3 is involved in multiple signalling pathways²³, raising the question of how its activity is selectively regulated by individual pathways. An emerging theme is that GSK3 is present, together with its substrates, in distinct complexes that are regulated by different upstream stimuli. Future study will determine whether Sgg/GSK3 forms a complex with Cos2 or Ci and whether Hh regulates Sgg/GSK3 within the complex. In vertebrates, three Gli proteins (Gli1, Gli2 and Gli3) are implicated in transducing Hh signals²⁶. Interestingly, all three Gli proteins contain multiple GSK3-phosphorylation consensus sites adjacent to PKA sites (our unpublished observation), raising the possibility that GSK3 may regulate Gli proteins in vertebrate Hh pathways. Hh and Wnt signalling pathways act in synergy in certain developmental contexts²⁷. Our finding that GSK3 is involved in both Hh and Wnt pathways raises the possibility that these two pathways might converge at GSK3 in certain developmental processes. □

Methods

Mutations and transgenes

sgg^{D127}, *Su(fu)*^{LP} and *smo*³ have been described^{7,28,29}. *Dp(1;2)w+70h* and *Dp(1;3)w+67K* are duplications of the *sgg* locus on the left arm of chromosome II and right arm of chromosome III, respectively (according to FlyBase, <http://flybase.bio.indiana.edu>). *UAS-mc**, *UAS-cos2*, *UAS-Ci-3P* and *UAS-Δarm* have been described^{6,15,21}. *hsp-flp.1*, *hsp-CD2* and *hsp-Myc-GFP*, as well as *dpp-lacZ*, *ptc-lacZ*, *wg-lacZ* and *hh-lacZ* reporter genes, have been described^{4,14,18,30}. Gal4 driver lines *MS1096* and *actin5c > CD2 > Gal4* have been described⁶. Cim1, Cim2 and Cim3 were generated by PCR-based site-directed

mutagenesis, and were subcloned into pUAST vector to generate *UAS-HA-Cim1*, *UAS-HA-Cim2* and *UAS-HA-Cim3* as described⁶. A dominant negative *Xenopus* GSK3 (ref. 13) was cloned into the pUAST vector to generate *UAS-DN-GSK3*.

Generating clones of marked mutant cells

Clones of mutant cells were generated by FLP/FRT-mediated mitotic recombination as described¹⁴. Genotypes for generating clones are as follows. *sgg*⁻ clones: *y sgg*^{D127}*hsp-flp.1/Y; dpp-lacZ(wg-lacZ or ptc-lacZ)/+; FRT82B hsp-CD2, y⁺ Dp(1;3)w+67K, sgg⁺/FRT82B hsp-Myc-GFP. sgg Su(fu) double-mutant clones: *y sgg*^{D127}*hsp-flp.1/Y; dpp-lacZ(wg-lacZ or ptc-lacZ)/+; FRT82B hsp-CD2, y⁺ Dp(1;3)w+67K, sgg⁺/FRT82B Su(fu)^{LP}. sgg smo double-mutant clones: *y sgg*^{D127}*hsp-flp.1/Y; Dp(1;2) sc¹⁹, y⁺ Dp(1;2)w+70h, sgg⁺ hsp-CD2, y⁺ FRT39E/smo³ stc FRT39E. smo⁻ clones expressing UAS-Ci transgenes: *y MS1096 hsp-flp.1/y w or Y; smo³ stc FRT39E/hsp-CD2, y⁺ FRT39E; UAS-HA-Ci(HA-Cim1, HA-Cim2 or HA-Cim3)/hh-lacZ*.***

Staining, culture, blotting and *in vitro* kinase assay

Standard protocols for immunofluorescence staining of imaginal discs were used³⁰. Antibodies used in this study, including anti-Ci (2A1), anti-βGal, anti-Sc, anti-CD2 and anti-Smo, have been described^{14,16,18,30}. Culture of cl-8 cells and drug treatment were performed as described⁵. Western blot analysis of protein extracts from wing imaginal discs was performed as described². For experiments involving LiCl treatment, imaginal discs were treated with 50 mM LiCl for up to 8 h before immunostaining, whereas cl-8 cells were treated with 50 mM LiCl for 4 h before lysis. For the *in vitro* kinase assay, GST–Ci fusion proteins were expressed in bacteria and purified using standard protocols, and were subjected to kinase assay with purified recombinant material of the PKA catalytic subunit and GSK3β (New England Biolabs) according to the supplier's protocols.

Received 8 February; accepted 7 March 2002.

Published online 24 March 2002, DOI 10.1038/nature733.

- Cadigan, K. M. & Nusse, R. Wnt signalling: a common theme in animal development. *Genes Dev.* **11**, 3286–3305 (1997).
- Aza-Blanc, P., Ramirez-Weber, F., Laget, M., Schwartz, C. & Kornberg, T. Proteolysis that is inhibited by Hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. *Cell* **89**, 1043–1053 (1997).
- Ohlmeier, J. T. & Kalderon, D. Hedgehog stimulates maturation of Cubitus interruptus into a labile transcriptional activator. *Nature* **396**, 749–753 (1998).
- Methot, N. & Basler, K. Hedgehog controls limb development by regulating the activities of distinct transcriptional activator and repressor forms of Cubitus interruptus. *Cell* **96**, 819–831 (1999).
- Chen, C. H. *et al.* Nuclear trafficking of Cubitus interruptus in the transcriptional regulation of Hedgehog target gene expression. *Cell* **98**, 305–316 (1999).
- Wang, G., Wang, B. & Jiang, J. Protein kinase A antagonizes Hedgehog signalling by regulating both the activator and repressor forms of Cubitus interruptus. *Genes Dev.* **13**, 2828–2837 (1999).
- Preat, T. Characterization of Suppressor of fused, a complete suppressor of the fused segment polarity gene of *Drosophila melanogaster*. *Genetics* **132**, 725–736 (1992).
- Ingham, P. W. Signaling by hedgehog family proteins in *Drosophila* and vertebrate development. *Curr. Opin. Genet. Dev.* **5**, 492–498 (1995).
- Villavicencio, E. H., Walterhouse, D. O. & Iannaccone, P. M. The sonic hedgehog-patched-gli pathway in human development and disease. *Am. J. Hum. Genet.* **67**, 1047–1054 (2000).
- Basler, K. & Struhl, G. Compartment boundaries and the control of *Drosophila* limb pattern by hedgehog protein. *Nature* **368**, 208–214 (1994).
- Tabata, T. & Kornberg, T. B. Hedgehog is a signalling protein with a key role in patterning *Drosophila* imaginal discs. *Cell* **76**, 89–102 (1994).
- Wang, Q. T. & Holmgren, R. A. The subcellular localization and activity of *Drosophila* cubitus interruptus are regulated at multiple levels. *Development* **126**, 5097–5106 (1999).
- Yost, C. *et al.* GBP, an inhibitor of GSK-3, is implicated in *Xenopus* development and oncogenesis. *Cell* **93**, 1031–1041 (1998).
- Jiang, J. & Struhl, G. Complementary and mutually exclusive activities of decapentaplegic and wingless organize axial patterning during *Drosophila* leg development. *Cell* **86**, 401–409 (1996).
- Zecca, M., Basler, K. & Struhl, G. Direct and long-range action of a Wingless morphogen gradient. *Cell* **87**, 833–844 (1996).
- Denef, N., Neubuser, D., Perez, L. & Cohen, S. M. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* **102**, 521–531 (2000).
- Price, M. A. & Kalderon, D. Proteolysis of cubitus interruptus in *Drosophila* requires phosphorylation by protein kinase A. *Development* **126**, 4331–4339 (1999).
- Jiang, J. & Struhl, G. Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb. *Nature* **391**, 493–496 (1998).
- Chen, Y., Cardinaux, J. R., Goodman, R. H. & Smolik, S. M. Mutants of cubitus interruptus that are independent of PKA regulation are independent of hedgehog signalling. *Development* **126**, 3607–3616 (1999).
- Sisson, J. C., Ho, K. S., Suyama, K. & Scott, M. P. Costal2, a novel kinesin-related protein in the Hedgehog signalling pathway. *Cell* **90**, 235–245 (1997).
- Wang, G., Amanai, K., Wang, B. & Jiang, J. Interactions with Costal2 and suppressor of fused regulate nuclear translocation and activity of cubitus interruptus. *Genes Dev.* **14**, 2893–2905 (2000).
- Klein, P. S. & Melton, D. A. A molecular mechanism for the effect of lithium on development. *Proc. Natl Acad. Sci. USA* **93**, 8455–8459 (1996).
- Kim, L. & Kimmel, A. R. GSK3, a master switch regulating cell-fate specification and tumorigenesis. *Curr. Opin. Genet. Dev.* **10**, 508–514 (2000).
- Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Nash, P. *et al.* Multisite phosphorylation of a CDK inhibitor sets a threshold for the onset of DNA replication. *Nature* **414**, 514–521 (2001).
- Ruiz i Altaba, A. Gli proteins and Hedgehog signalling: development and cancer. *Trends Genet.* **15**, 418–425 (1999).

27. Roelink, H. Tripartite signalling of pattern: interactions between Hedgehogs, BMPs and Wnts in the control of vertebrate development. *Curr. Opin. Neurobiol.* **6**, 33–40 (1995).
28. Ruel, L., Pantescio, Y., Lutz, P., Simpson, P. & Bourouis, M. Functional significance of a family of protein kinases encoded at the *shaggy* locus in *Drosophila*. *EMBO J.* **12**, 1657–1669 (1993).
29. Chen, Y. & Struhl, G. Dual roles for patched in sequestering and transducing Hedgehog. *Cell* **87**, 553–563 (1996).
30. Jiang, J. & Struhl, G. Protein kinase A and Hedgehog signalling in *Drosophila* limb development. *Cell* **80**, 563–572 (1995).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com>).

Acknowledgements

We thank G. Struhl, D. Kalderon, R. Holmgren, S. Cohen, T. Kornberg and M. Peifer for reagents; and G. Struhl, K. Wharton and L. Parada for comments. This work was supported by the National Institutes of Health. J.J. is a Searle Scholar supported by the Chicago Community Trust, a Eugene McDermott Endowed Scholar in Biomedical Research, and a Leukemia and Lymphoma Society Special Fellow.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.J. (e-mail: Jiang05@utsw.swmed.edu).

DNMT1 and DNMT3b cooperate to silence genes in human cancer cells

Ina Rhee^{*†‡}, Kurtis E. Bachman[†], Ben Ho Park^{*†}, Kam-Wing Jair[†], Ray-Whay Chiu Yen[†], Kornel E. Schuebel[†], Hengmi Cui^{§||}, Andrew P. Feinberg^{§||¶}, Christoph Lengauer[†], Kenneth W. Kinzler^{†‡}, Stephen B. Baylin^{†‡#} & Bert Vogelstein^{*†‡#}

^{*} The Howard Hughes Medical Institute, [†] The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, [‡] Program in Human Genetics, [§] Institute of Genetic Medicine, ^{||} Department of Medicine, [¶] Departments of Molecular Biology & Genetics and Oncology, and [#] Program in Cellular and Molecular Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21231, USA

Inactivation of tumour suppressor genes is central to the development of all common forms of human cancer¹. This inactivation often results from epigenetic silencing associated with hypermethylation rather than intragenic mutations^{2–7}. In human cells, the mechanisms underlying locus-specific or global methylation patterns remain unclear^{8,9}. The prototypic DNA methyltransferase, Dnmt1, accounts for most methylation in mouse cells^{10,11}, but human cancer cells lacking DNMT1 retain significant genomic methylation and associated gene silencing¹². We disrupted the human *DNMT3b* gene in a colorectal cancer cell line. This deletion reduced global DNA methylation by less than 3%. Surprisingly, however, genetic disruption of both *DNMT1* and *DNMT3b* nearly eliminated methyltransferase activity, and reduced genomic DNA methylation by greater than 95%. These marked changes resulted in demethylation of repeated sequences, loss of insulin-like growth factor II (*IGF2*) imprinting, abrogation of silencing of the tumour suppressor gene *p16^{INK4a}*, and growth suppression. Here we demonstrate that two enzymes cooperatively maintain DNA methylation and gene silencing in human cancer cells, and provide compelling evidence that such methylation is essential for optimal neoplastic proliferation.

Homologous recombination was used to disrupt both of the alleles of *DNMT3b* in the colorectal carcinoma line, HCT116. We chose HCT116 because it exhibits hypermethylation-associated

gene silencing, a diploid karyotype, and susceptibility to targeted homologous integration¹². Successful targeting replaced exons 2 to 21 of *DNMT3b* with either a promoterless neomycin- or hygromycin-resistance gene expressed from the endogenous *DNMT3b* promoter (Fig. 1a). This deletion encompasses the active site prolylcysteiny dipeptide, which is critical for catalysis¹³. Mutant genotypes and loss of the *DNMT3b* gene product were confirmed by Southern and western blotting (Fig. 1b, c). Two *DNMT3b*^{−/−} clones were tested in detail and found to contain greater than 87% of wild-type DNA methyltransferase activity using the standard synthetic substrate poly(dI-dC)·poly(dI-dC) (Fig. 1d). Consistent with this small change in methyltransferase activity, *DNMT3b*^{−/−} cells retained greater than 97% of wild-type genomic 5-methylcytosine (m⁵C) content (Fig. 2a), as quantified by reversed-phase high performance liquid chromatography (HPLC)^{14,15}. Modest demethylation of pericentromeric satellite 2 sequences was observed in the *DNMT3b*^{−/−} clones (Fig. 3a), but no changes in the characteristic methylation or expression of single-copy genes, including *p16^{INK4a}*, were observed (Fig. 4).

We next pursued disruption of both *DNMT3b* and *DNMT1* in the same cells. *LoxP* sites present in the *DNMT3b* targeting construct facilitated the serial targeting of these genes. Expression of Cre recombinase resulted in the removal of the neomycin- and hygromycin-resistance genes from the targeted *DNMT3b* alleles, allowing us to introduce a previously described hygromycin targeting vector¹² into the *DNMT1* locus. Subsequent targeting with a neomycin-based *DNMT1* targeting vector yielded cells bearing homozygous deletions of both *DNMT3b* and *DNMT1*. We obtained a total of eight independent double knockout (DKO) clones with this genotype, as confirmed by Southern and western blots (Fig. 1b, c). DKO 1 and DKO 2 are representative of seven clones (DKO 1–7) that behaved identically.

DNA methyltransferase activity was nearly abolished in DKO 1 and DKO 2 cells, whereas it was only reduced by approximately 93% in *DNMT1*^{−/−} cells (Fig. 1d). Similarly, DKO cells exhibited marked loss of genomic m⁵C (Fig. 2a). Instead of the expected ~23% decrease in m⁵C content, representing the sum of the approximately 20% (ref. 12) and 3% (Fig. 2a) reductions attributable to *DNMT1* and *DNMT3b* disruptions, respectively, we observed a roughly 95% reduction in m⁵C content in the representative clones (Fig. 2a). *HpaII* digests of genomic DNA corroborated this reduction in m⁵C content (Fig. 2b). Cleavage by this m⁵C-sensitive restriction endonuclease appeared virtually complete in the DKO clones, as it was comparable to that achieved after treatment with 5-aza-2'-deoxycytidine (5-aza-dC), a pharmacological inhibitor of DNA methylation, or digestion with *MspI*, an isoschizomer of *HpaII* that is insensitive to m⁵C. In contrast, the digestion patterns of genomic DNA from cells with deletions of either *DNMT1* or *DNMT3b* alone were very similar to those of parental cells (Fig. 2b).

It was possible that differences between the *DNMT1*^{−/−} lines and DKO lines were due to slight differences at the *DNMT1* locus resulting from the distinct targeting strategies used. Specifically, the generation of *DNMT1*^{−/−} cells involved Cre-mediated removal of the antibiotic resistance cassette from one allele before targeting of the second allele, whereas DKO cells retained antibiotic resistance cassettes in both alleles. Two experiments addressed this concern. First, new *DNMT1*^{−/−} cells were created by targeted deletion of both *DNMT1* alleles without using Cre to remove the antibiotic resistance genes. Second, the antibiotic resistance cassettes were deleted from DKO cells through Cre expression. Neither manipulation resulted in a change of methyltransferase activity or global methylation levels (Figs 1d and 2a). Accordingly, the growth characteristics and methylation patterns of these new lines were identical to those of the original clones (not shown).

The global loss of DNA methylation in the DKO cells was associated with distinctly abnormal methylation of each endogenous locus analysed. Southern blot analyses of genomic DNA after

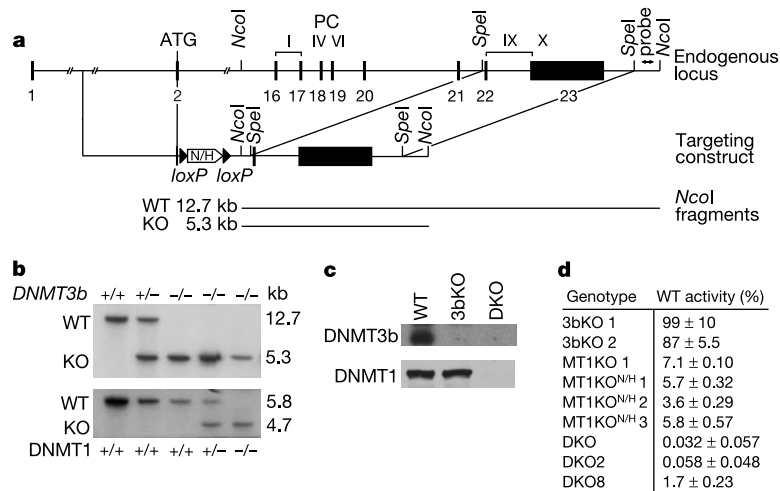


Figure 1 Human cells lacking *DNMT3b* and *DNMT1*. **a**, **b**, Targeted deletion of *DNMT3b*. Numbered boxes represent exons. Neomycin- or hygromycin-resistance genes (N/H) replace conserved DNA cytosine methyltransferase motifs (Roman numerals), including the critical prolylcysteine dipeptide (PC). A probe distinguishing *NcoI* fragments from wild-type (WT) and targeted (KO) alleles allowed Southern blot confirmation (**b**) of the indicated

genotypes. **c**, Western blot of wild-type (WT), *DNMT3b*^{-/-} (3bKO) and *DNMT3b*^{-/-}; *DNMT1*^{-/-} (DKO) cells using antibodies described in Methods. **d**, DNA methyltransferase activity as a percentage of wild-type activity (± s.d.) of triplicate measurements. MT1KO^{N/H} designates *DNMT1*^{-/-} (MT1KO) clones retaining two drug cassettes (see text).

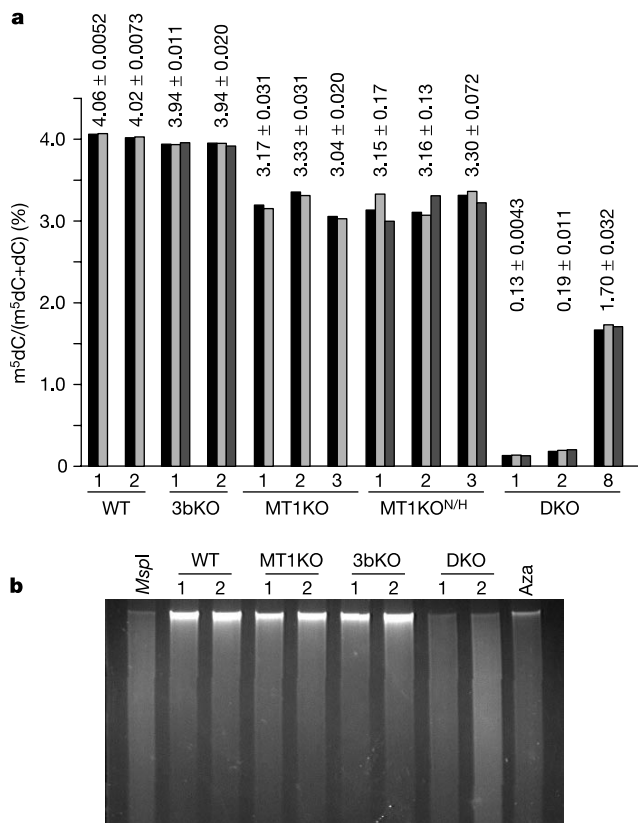


Figure 2 *DNMT3b*^{-/-}; *DNMT1*^{-/-} cells are demethylated. **a**, Independent measurements of 5-methylcytosine by reversed-phase HPLC separation of genomic DNA digests from clones of the indicated genotypes, designated as in Fig. 1. Results are expressed as a percentage of the total cytosine pool (± s.d.) from duplicate or triplicate measurements. **b**, Genomic DNA from clones of the indicated genotype or wild-type (WT) cells treated with 5-aza-dC (Aza) was digested with *HpaII*, which does not cleave methylated sites. Comparison with wild-type DNA digested with the isoschizomer *MspI*, which is insensitive to methylation, reveals extensive demethylation in the DKO clones.

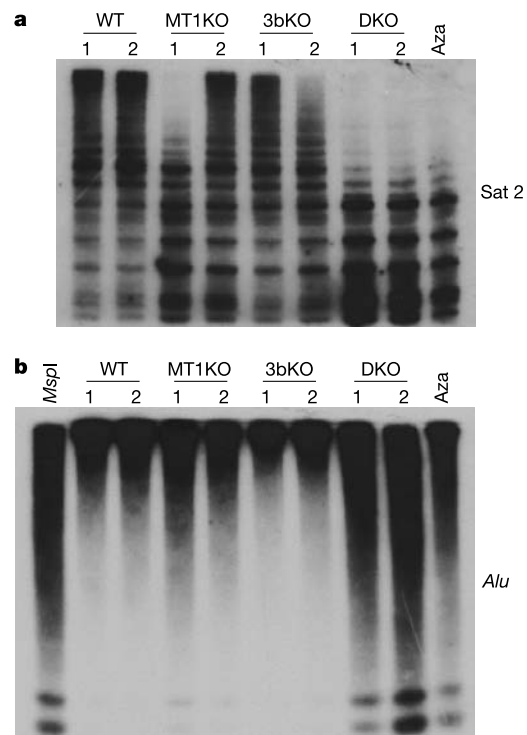


Figure 3 Repetitive regions are demethylated in *DNMT3b*^{-/-}; *DNMT1*^{-/-} cells. **a**, **b**, Southern blotting was carried out with genomic DNA from independent clones of the indicated genotype (as designated in Fig. 1) or from wild-type (WT) cells treated with 5-aza-dC (Aza). DNA was digested with *BstBI* (**a**) or *HpaII* (**b**), which do not cleave methylated sites, or with *MspI* (**b**), which is insensitive to methylation. Oligonucleotide probes were derived from consensus sequences for juxta-centromeric satellite 2 (Sat 2) sequences (**a**) or *Alu* repeats (**b**).

digestion with m^5C -sensitive restriction endonucleases revealed marked loss of methylation at classical satellite 2 sequences (Fig. 3a), as well as at the more widely distributed *Alu* family of repeated sequences (Fig. 3b). The patterns observed in the DKO cells were similar to those observed after 5-aza-dC treatment (Fig. 3a, b) or after digestion with *MspI* (Fig. 3b).

TIMP-3 is a single-copy gene that is silenced and hypermethylated in diverse human cancer cells¹⁶. Promoter methylation was evaluated through methylation-specific polymerase chain reaction (MSP)¹⁷, which entails treatment of genomic DNA with sodium bisulphite to allow the design of PCR primers that distinguish methylated and unmethylated sequences. The *TIMP-3* promoter was fully methylated in parental HCT116 cells and fully unmethylated in peripheral blood lymphocytes, in which *TIMP-3* is

expressed. Significant promoter demethylation was observed in the DKO cells, comparable to demethylation observed in parental cells treated with 5-aza-dC (Fig. 4a). This demethylation was associated with expression of the *TIMP-3* transcript (Fig. 4b).

The normal imprinting of *IGF2* is disrupted in a variety of human tumours^{18–21}. An intragenic restriction polymorphism for which HCT116 cells are heterozygous²² allowed relative quantification of *IGF2* alleles, and a loss of imprinting (LOI) index was derived for each cell line²². In single mutant *DNMT1*^{-/-} or *DNMT3b*^{-/-} cells, as in wild-type cells, allele A was expressed considerably more robustly than the silenced allele B. LOI indices of 17–32% were within the range of normally imprinted cells. However, DKO cells expressed both alleles at equivalent levels, with LOI indices of 95–109%, representing a complete loss of imprinting (Fig. 4c).

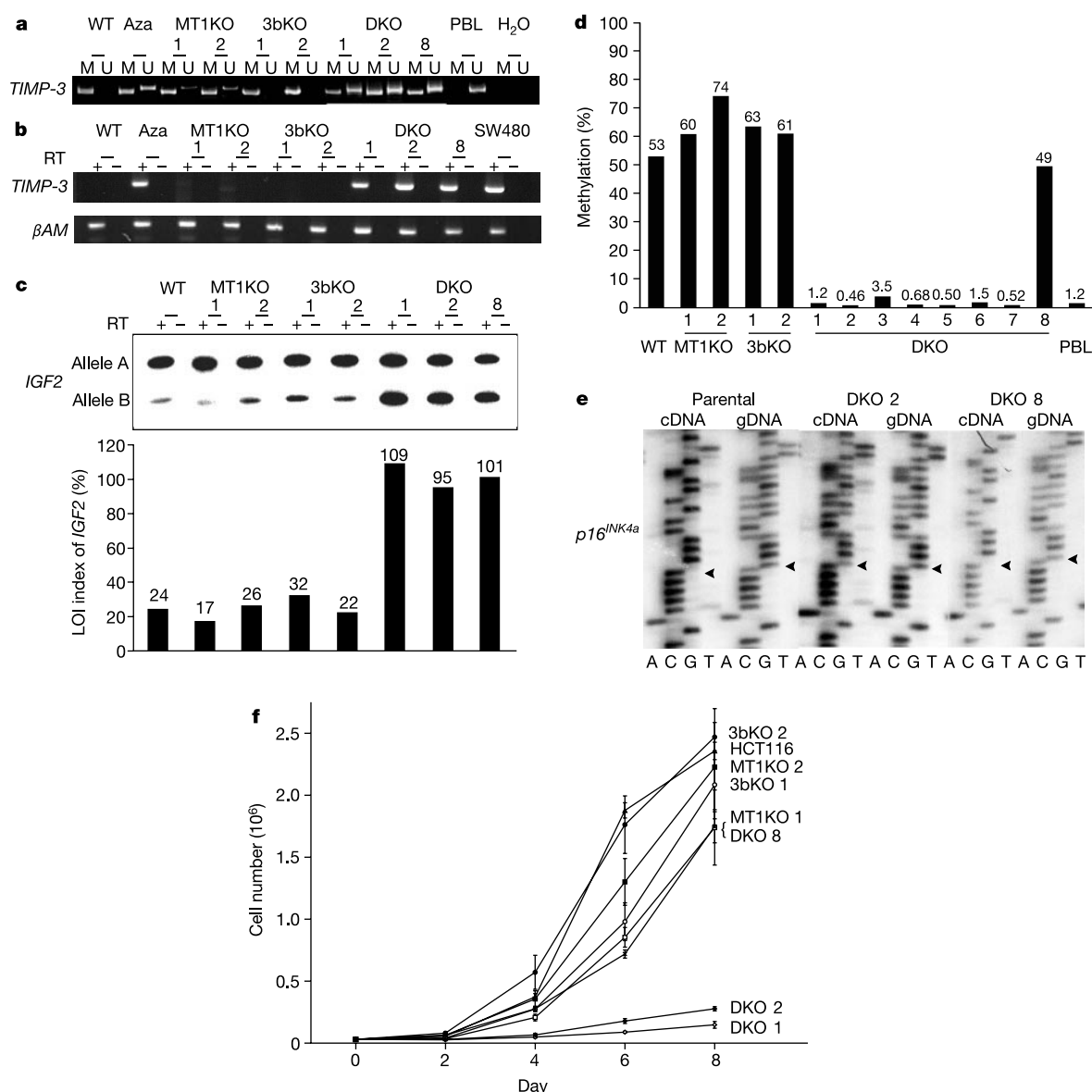


Figure 4 Gene silencing in *DNMT3b*^{-/-};*DNMT1*^{-/-} cells. **a**, MSP of *TIMP-3*. HCT116 cells (WT) and peripheral blood lymphocytes (PBL) harbour exclusively methylated (M) or unmethylated (U) alleles, respectively. **b**, RT-PCR of *TIMP-3*, which is expressed in SW480 cells. β -amyloid (β AM) was used as a control. **c**, *Apal* digestion of *IGF2* cDNA amplification products to determine loss of imprinting (LOI), as described²². The bars in

the histogram correspond to the cell types shown at top of the panel. **d**, *p16*^{INK4a} MSP to quantify the percentage of alleles that are methylated. **e**, Sequencing of *p16*^{INK4a} amplification products to delineate frameshifted (arrow) and wild-type allele expression. gDNA, genomic DNA. **f**, Growth curves depicting triplicate cell counts with standard deviations at two-day intervals.

The $p16^{INK4a}$ gene exemplifies tumour suppressor genes whose silencing is dependent on methylation⁶. In HCT116 cells, one allele of $p16^{INK4a}$ is wild type and the other harbours a truncating frameshift mutation. The promoter of the wild-type allele is exclusively methylated in these cells, and only the mutant allele is expressed²³. Quantitative MSP revealed that methylation of the $p16^{INK4a}$ promoter was retained in cells in which either *DNMT1* or *DNMT3b* was individually disrupted, but reduced by greater than 98% in DKO clones 1 to 7 (Fig. 4d). Importantly, demethylation of the $p16^{INK4a}$ promoter was associated with expression of the wild-type $p16^{INK4a}$ allele (DKO 2 is a representative example shown in Fig. 4e). No such expression of the wild-type allele was detected in parental cells or in single mutant derivatives (Fig. 4e and data not shown).

Our studies demonstrate that DNMT3b and DNMT1 cooperatively maintain virtually all methylation in HCT116 cells. This includes the normal methylation of repeated sequences and the silencing-associated methylation characteristic of genes like *IGF2*, *TIMP-3* and $p16^{INK4a}$. The marked demethylation of double mutant clones compared with individual *DNMT1* and *DNMT3b* mutants suggests an unsuspected degree of functional redundancy among enzymes previously characterized as having predominantly 'de novo' or 'maintenance' functions^{11,24,25}.

These results also establish that methylation is essential for silencing tumour suppressor genes in human cancer cells. Although this idea was supported by experiments using 5-aza-dC, this drug is toxic and has multiple effects on RNA and DNA metabolism^{26,27}. We showed that specific targeting of methyltransferase genes resulted in re-expression of the silenced wild-type $p16^{INK4a}$ allele to the level of the unmethylated mutant $p16^{INK4a}$ allele. Moreover, our data illustrate that DNMT1/DNMT3b-dependent silencing of tumour suppressor genes is essential for growth control. Seven out of eight DKO clones suffered pronounced growth suppression when epigenetic silencing was disrupted through targeted disruption of *DNMT1* and *DNMT3b*. In contrast, one aberrant clone, DKO 8, exhibited higher global m⁵C content (Fig. 2a) and retained both $p16^{INK4a}$ silencing (Fig. 4d, e) and robust growth (Fig. 4f). The DKO 8 clone demonstrates that tumour cells can use enzymes other than DNMT1 and DNMT3b to maintain methylation of critical sites under certain circumstances.

The relationship between DNA methylation and growth is further supported by recent studies showing that loss of the *Dnmt1* gene in mouse fibroblasts was associated with re-expression of a silenced *p21^{WAF1/CIP1}* gene and p53-dependent cell death²⁸. In contrast, the DKO clones described here were fully compatible with continued clonal growth despite an intact p53 pathway. It will be of interest to determine whether these results apply to other human cell types. Notably, DKO cells efficiently silenced exogenous retroviral genes and remained sensitive to 5-aza-dC-mediated toxicity (Supplementary Information Figs 1 and 2). In addition to DNMT1 and DNMT3b, the only other known functional human methyltransferase is DNMT3a^{11,13}. *DNMT3a* was expressed at similar, low levels in all DKO lines (including DKO 8), as assessed by reverse transcriptase (RT)-PCR analysis (not shown). Thus, a new enzyme may be responsible for residual methylation in DKO cells. These lines represent a unique resource for analysing the precise biochemical properties of known as well as postulated methyltransferase enzymes and the relationship between DNA methylation and specific physiological and biochemical properties of mammalian cells. □

Methods

Targeting vectors

The *DNMT1* targeting construct has been described previously¹². An otherwise identical vector, containing a neomycin-resistance gene in place of the hygromycin-resistance gene, was also constructed. A genomic clone containing *DNMT3b* was obtained by screening a human bacterial artificial chromosome (BAC) library (Research Genetics). A 3.0-kilobase (kb) fragment upstream of exon 2 and a 4.5-kb *SpeI* fragment downstream of exon 21 were

incorporated on either side of a promoterless neomycin- or hygromycin-resistance gene flanked by *loxP* sites. This strategy allowed Cre-mediated excision of both drug markers after serial targeting of both of the *DNMT3b* alleles, and thus the subsequent targeting of *DNMT1*, using both the hygromycin- and neomycin-based vectors, to create the DKO clones. A PCR screen for targeted alleles was developed as described²⁹. Further details about the targeting vector and screening PCR primers are available from the authors on request.

Cell culture, transfection and screening for recombinants

HCT116 cells (American Type Culture Collection) were cultured in McCoy's 5A modified medium supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. HCT116 cells were transfected with *XmaI*-linearized *DNMT3b* targeting vector using LipofectAmine (Life Technologies), and selected in growth medium supplemented with either 0.4 mg ml⁻¹ geneticin (Life Technologies) or 0.05 mg ml⁻¹ hygromycin (Calbiochem). Homologous recombinants identified by PCR were confirmed by Southern blot (Fig. 1). A *DNMT3b*^{-/-} clone was infected with an adenovirus expressing Cre recombinase to yield drug-sensitive clones, and one such clone was subjected to deletion of DNMT1. We studied two independent *DNMT3b*^{-/-} and eight independent DKO clones.

Western blot analysis and DNA methyltransferase assays

A rabbit polyclonal antibody (QCB/BioSource International) was raised against a fusion protein containing residues 376–390 (ENKTRRRRTADDSSATS) of DNMT3b1 (GenBank accession no. AAD53063). The polyclonal antibody HMT1509 recognizing DNMT1 has been described³⁰. Nuclear extracts for western blotting were prepared using the NE-PER Nuclear and Cytoplasmic Extraction Reagents Kit (Pierce), according to the manufacturer's protocol. A total of 50 µg of extract was resolved by 7% SDS–polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes (Bio-Rad). Immunocomplexes were detected with a horseradish-peroxidase-conjugated donkey anti-rabbit secondary antibody (ASK) (Amersham Pharmacia Biotech) and visualized using the SuperSignal chemiluminescence kit according to the manufacturer's protocol (Pierce). DNA methyltransferase enzyme activity was measured as described¹² using 15 µg of protein lysate, 3 µCi S-adenosyl-L-[methyl-³H] methionine (Amersham Pharmacia Biotech), and 0.5 µg poly(dI-dC)-poly(dI-dC) (Amersham Pharmacia Biotech). Results from parallel reactions carried out without DNA were subtracted from experimental values.

High-performance liquid chromatography

Approximately 40 µg of RNA-free genomic DNA, prepared using the Blood and Cell Culture DNA Midi Kit (Qiagen), was quantitatively digested with nuclease P1 (Roche Molecular Biochemicals) and calf intestinal alkaline phosphatase (Sigma), as described¹⁴. Samples were separated essentially as described¹⁵ on a reversed-phase column (Supelcosil LC-18 DB) at room temperature, monitoring absorbances at 275 nm and 285 nm. Peak assignments were confirmed using deoxyribonucleoside standards (Sigma). We expressed 5-methylcytosine content as a percentage of the total cytosine pool, using peak areas after correction for extinction coefficients.

5-Aza-dC treatment and methylation-sensitive Southern blotting

Growth medium supplemented with 5 µM 5-aza-dC (Sigma) was added to log-phase cells every 24 h over a 72-h period before isolating genomic DNA (QiaAmp DNA Blood Mini Kit, Qiagen) or RNA (RNagents, Promega). DNA from treated and untreated cells was digested with *HpaII*, *MspI* or *BstBI* (New England Biolabs), and resolved on 1.2% or 2.5% agarose gels for classical satellite and *Alu* probes, respectively. Southern blotting was performed according to standard methods, using probes for satellite 2 (S2) and *Alu* family (*Alu3*) sequences¹².

Analysis of gene expression and methylation

Approximately 1 µg of genomic DNA was treated with sodium bisulphite and subjected to MSP as described¹⁷. MSP primer sequences that specifically recognize unmethylated *TIMP-3* alleles were 5'-TTGTTGTGTTTATTGTTTGTGTT-3' and 5'-ATTACCATA CACACCACCAACA-3', and the corresponding primers that recognize methylated alleles were 5'-CGTTGCGTTTATTTCGTTTCGTC-3' and 5'-TACGCGCCGCCGACG-3'. *TIMP-3* was amplified from complementary DNA using the primers 5'-GCTGTGCAAC TTCGTGGAGAGG-3' and 5'-CTCGGTACACAGTCGAGTAGCC-3'. Control amplifications were performed using primers 5'-GTGAAGATGGATGAGAATTCGG-3' and 5'-AAAGAACTGTAGGTTGGATTTCG-3' against the β-amyloid gene. MSP primers that recognize unmethylated or methylated $p16^{INK4a}$ alleles have been described¹⁷. A 1:50,000 dilution of the intercalating dye SYBR green (Molecular Probes) was added to the reactions, and amplification products were quantified in real-time using the iCycler Thermal Cycler (Bio-Rad). The number of amplification cycles required to reach a statistical threshold using primers specific for methylated alleles was subtracted from the corresponding number derived for the primers specific for unmethylated alleles. This cycle number difference (*d*) was converted to the percentage of $p16^{INK4a}$ alleles that are methylated using the formula: methylation = $2^d / (2^d + 1)$. We amplified and sequenced the region of $p16^{INK4a}$ containing the frameshift mutation as described¹². Hot-stop PCR²² was used to assess imprinting of the *IGF2* locus, which contains a transcribed *Apal* polymorphism that allows distinction of the two alleles, A and B. Briefly, PCR amplification was followed by a single primer-extension step using ³²P-labelled primers. Products were digested with *Apal* and quantified after electrophoresis using a PhosphorImager (Molecular Dynamics). We determined the loss of imprinting (LOI) index by the ratio of (allele B × 100)/allele A.

In vitro growth assay

Cells were seeded in triplicate in 12-well plates at a concentration of 3×10^4 cells per well. Cells were collected at days 2, 4, 6 and 8, and viable cells, as assessed by trypan blue exclusion, were counted in a haemocytometer.

Received 9 October 2001; accepted 14 February 2002.

- Vogelstein, B. & Kinzler, K. W. *The Genetic Basis of Human Cancer* (McGraw-Hill Health Professions Division, New York, 1998).
- Siegrfried, Z. & Cedar, H. DNA methylation: a molecular lock. *Curr. Biol.* **7**, R305–R307 (1997).
- Bird, A. P. & Wolffe, A. P. Methylation-induced repression—belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
- Robertson, K. D. & Jones, P. A. DNA methylation: past, present and future directions. *Carcinogenesis* **21**, 461–467 (2000).
- Tycko, B. Epigenetic gene silencing in cancer. *J. Clin. Invest.* **105**, 401–407 (2000).
- Baylin, S. B. & Herman, J. G. DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet.* **16**, 168–174 (2000).
- Ponder, B. A. Cancer genetics. *Nature* **411**, 336–341 (2001).
- Eads, C. A. et al. CpG island hypermethylation in human colorectal tumors is not associated with DNA methyltransferase overexpression. *Cancer Res.* **59**, 2302–2306 (1999).
- Schmutte, C., Yang, A. S., Nguyen, T. T., Beart, R. W. & Jones, P. A. Mechanisms for the involvement of DNA methylation in colon carcinogenesis. *Cancer Res.* **56**, 2375–2381 (1996).
- Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
- Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247–257 (1999).
- Rhee, I. et al. CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* **404**, 1003–1007 (2000).
- Bestor, T. H. The DNA methyltransferases of mammals. *Hum. Mol. Genet.* **9**, 2395–2402 (2000).
- Kuo, K. C., McCune, R. A., Gehrke, C. W., Midgett, R. & Ehrlich, M. Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res.* **8**, 4763–4776 (1980).
- Feinberg, A. P., Gehrke, C. W., Kuo, K. C. & Ehrlich, M. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res.* **48**, 1159–1161 (1988).
- Bachman, K. E. et al. Methylation-associated silencing of the tissue inhibitor of metalloproteinase-3 gene suggest a suppressor role in kidney, brain, and other human cancers. *Cancer Res.* **59**, 798–802 (1999).
- Herman, J. G., Graff, J. R., Myohanen, S., Nelkin, B. D. & Baylin, S. B. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc. Natl Acad. Sci. USA* **93**, 9821–9826 (1996).
- Rainier, S. et al. Relaxation of imprinted genes in human cancer. *Nature* **362**, 747–749 (1993).
- Ogawa, O. et al. Constitutional relaxation of insulin-like growth factor II gene imprinting associated with Wilms' tumour and gigantism. *Nature Genet.* **5**, 408–412 (1993).
- Steenman, M. et al. Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nature Genet.* **7**, 433–439 (1994).
- Cui, H., Horon, I. L., Ohlsson, R., Hamilton, S. R. & Feinberg, A. P. Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability. *Nature Med.* **4**, 1276–1280 (1998).
- Uejima, H., Lee, M. P., Cui, H. & Feinberg, A. P. Hot-stop PCR: a simple and general assay for linear quantitation of allele ratios. *Nature Genet.* **25**, 375–376 (2000).
- Myohanen, S. K., Baylin, S. B. & Herman, J. G. Hypermethylation can selectively silence individual p16INK4A alleles in neoplasia. *Cancer Res.* **58**, 591–593 (1998).
- Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genet.* **19**, 219–220 (1998).
- Lyko, F. et al. Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in *Drosophila*. *Nature Genet.* **23**, 363–366 (1999).
- Santi, D. V., Garrett, C. E. & Barr, P. J. On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell* **33**, 9–10 (1983).
- Juttermann, R., Li, E. & Jaenisch, R. Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc. Natl Acad. Sci. USA* **91**, 11797–11801 (1994).
- Jackson-Grusby, L. et al. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nature Genet.* **27**, 31–39 (2001).
- Chan, T. A., Hermeking, H., Lengauer, C., Kinzler, K. W. & Vogelstein, B. 14-3-3 σ is required to prevent mitotic catastrophe after DNA damage. *Nature* **401**, 616–620 (1999).
- Vertino, P. M., Yen, R. W., Gao, J. & Baylin, S. B. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5)-methyltransferase. *Mol. Cell Biol.* **16**, 4555–4565 (1996).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank S. R. Lee and S. G. Rhee for assistance with the HPLC analysis. This work was supported by the Clayton Fund, the V Foundation, and by grants from the National Institutes of Health.

Competing interests statement

The authors declare competing financial interests: details accompany the paper on Nature's website (<http://www.nature.com>).

Correspondence and requests for materials should be addressed to S.B.B. (e-mail: sbaylin@jhmi.edu) or B.V. (e-mail: vogelbe@welch.jhu.edu).

Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase

James P. Jackson, Anders M. Lindroth, Xiaofeng Cao & Steven E. Jacobsen

Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, California 90095, USA

Gene silencing in eukaryotes is associated with the formation of heterochromatin, a complex of proteins and DNA that block transcription. Heterochromatin is characterized by the methylation of cytosine nucleotides of the DNA, the methylation of histone H3 at lysine 9 (H3 Lys 9), and the specific binding of heterochromatin protein 1 (HP1) to methylated H3 Lys 9 (refs 1–7). Although the relationship between these chromatin modifications is generally unknown, in the fungus *Neurospora crassa*, DNA methylation acts genetically downstream of H3 Lys 9 methylation⁸. Here we report the isolation of KRYPTONITE, a methyltransferase gene specific to H3 Lys 9, identified in a mutant screen for suppressors of gene silencing at the *Arabidopsis thaliana* SUPERMAN (SUP) locus. Loss-of-function kryptonite alleles resemble mutants in the DNA methyltransferase gene CHROMOMETHYLASE3 (CMT3)⁹, showing loss of cytosine methylation at sites of CpNpG trinucleotides (where N is A, C, G or T) and reactivation of endogenous retrotransposon sequences. We show that CMT3 interacts with an *Arabidopsis* homologue of HP1, which in turn interacts with methylated histones. These data suggest that CpNpG DNA methylation is controlled by histone H3 Lys 9 methylation, through interaction of CMT3 with methylated chromatin.

Heterochromatin contains a characteristic set of post-translational histone modifications including H3 Lys 9 methylation, which is carried out by the SET domains of the Su(var)3-9 class proteins^{1,2}. Genetic studies show that Su(var)3-9 proteins are essential for proper assembly of heterochromatin. For example, *Drosophila* su(var)3-9 mutants suppress a silencing phenomenon called position effect variegation (PEV)¹⁰, and *clr4* mutants de-repress silent mating-type loci in *Schizosaccharomyces pombe*^{11,12}. Mice lacking SUV39h genes show defects in pericentromeric heterochromatin, chromosome instabilities, and increased tumorigenesis¹³. The 'histone code' hypothesis^{3,4} proposes that histone modifications direct the binding of specific proteins that mediate chromatin function. For instance, the chromo domain of HP1 has been shown to specifically bind methylated H3 Lys 9, and this binding is essential for heterochromatin formation *in vivo*^{2,5–7}. Here we provide evidence that a histone code influences the enzymes that methylate DNA.

We performed a screen for ethylmethane sulphonate (EMS)-induced suppressors of an epigenetic allele of the SUP locus (the *clark kent-st* allele, *clk-st*)^{9,14}. *clk-st* plants show defects in the number of floral organs (Fig. 1a) owing to cytosine methylation and silencing of the SUP gene. In *clk-st*, SUP is methylated not only at CpG dinucleotides but also at CpNpG sites and asymmetric sites (cytosines not present in CpG or CpNpG contexts). We previously reported that nine *clk-st* suppressor mutants were loss-of-function alleles of the DNA methyltransferase CHROMOMETHYLASE3 (CMT3)⁹. *cmt3* mutants reduce CpNpG DNA methylation and reactivate the expression of SUP, PAI2 and a subset of retrotransposons^{9,15}. Further analysis of the *clk-st* suppressor mutants identified three alleles of a new locus, which we have named KRYPTONITE (KYP) (Fig. 1a). These mutants, *kyp-1* to *kyp-3*, are recessive and show similar phenotypes, suggesting that they are loss-of-function mutants. The *kyp-2* mutant also suppresses a different

clark kent allele, *clk-3* (ref. 14). Other than suppression of the *clk* phenotype, the *kyp* mutants did not exhibit morphological defects even after extensive inbreeding.

The *KYP* gene was cloned and found to code for a protein of 624 amino acids containing a SET domain (Fig. 1b, c). A phylogenetic study of *Arabidopsis* SET proteins suggests that *KYP* is most similar to the Su(var)3-9 class of histone H3 Lys 9 methyltransferases¹⁶. This study described nine related *Arabidopsis* sequences listed as SU(VaR)3-9 homologues 1–9, with *KYP* listed as number 4. Alignment of *KYP* with several H3 methyltransferases shows conservation of sequences critical for *in vitro* methylase activity including the cysteine-rich pre-SET and post-SET motifs and specific residues within the SET domain (Fig. 1c)^{1,2}.

To test whether *KYP* methylates histones, we expressed a fusion protein of glutathione S-transferase (GST) and the *KYP* SET domain, and performed *in vitro* methylation assays. Similar to a

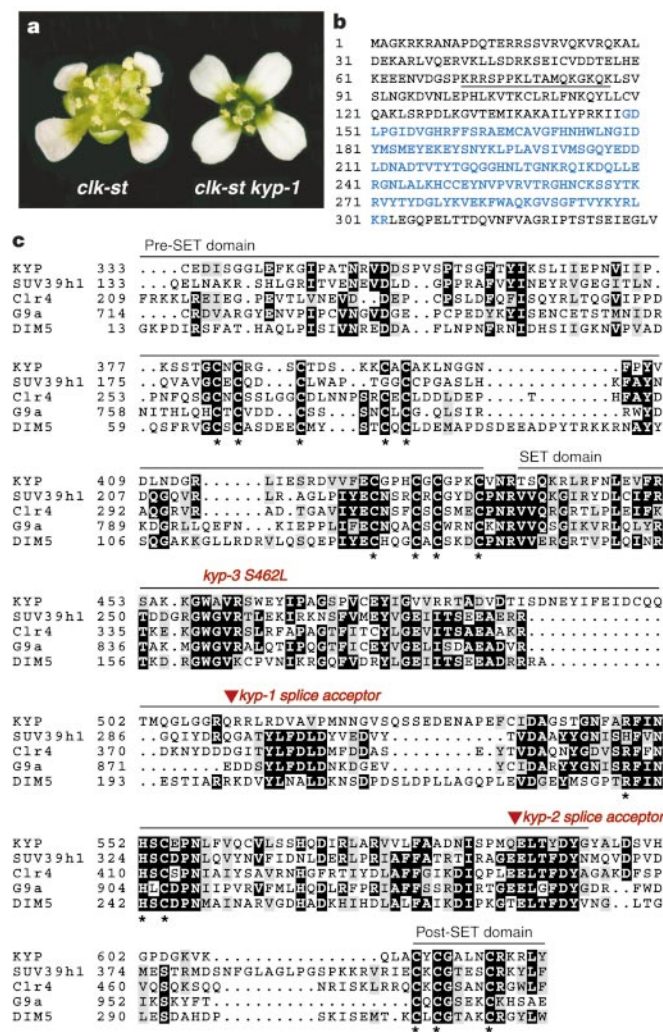


Figure 1 KRYPTONITE mutants. **a**, A *clk-st* flower containing nine stamens and a defective gynoecium and a *kyp-1*-suppressor mutant flower showing the normal six stamens and a normal gynoecium. **b**, KYP amino acids 1–332, showing a nuclear localization signal (underlined) and the YDG domain (blue)¹⁶. **c**, KYP amino acids 333–624 constituting the pre-SET, SET and post-SET domains, aligned with four histone H3 methyltransferases, human SUV39h1 (NP_003164), *Schizosaccharomyces pombe* Clr4 (T43745), human G9a (NP_006700) and *Neurospora* DIM5 (AF419248). Asterisks in the pre-SET and post-SET domains mark conserved cysteines and those in the SET domain mark residues important for methylase activity¹. Triangles mark splice junctions affected in *kyp-1* and *kyp-2*.

GST–SUV39h1 control, GST–*KYP* methylated histone H3 (Fig. 2a). GST–*KYP* did not methylate histones H1, H2A, H2B or H4. *KYP* also methylated a GST fusion protein containing the 57 amino-terminal amino acids of histone H3, but not a mutant fusion protein in which Lys 9 was mutated to arginine (Fig. 2b)¹⁷. Thus, like other Su(var)3-9 class proteins, *KYP* is a H3 Lys 9 methyltransferase. All three *kyp* alleles are predicted to reduce or eliminate the function of the SET domain (Fig. 1c), suggesting that H3 Lys 9 methyltransferase activity is critical for *KYP* function.

We determined the effect of *kyp* on DNA methylation using both bisulphite genomic sequencing and Southern blot analysis using methylation-sensitive restriction enzymes. A bisulphite sequencing analysis of the *SUP* gene, comparing the methylation profiles of *kyp-1* with those of the *cmt3-7* and *met1* mutants (*met1* is a recessive allele of the DNMT1-like MET1 CpG methyltransferase^{18–20}), is shown in Fig. 3a. In the *kyp-1* mutant, *SUP* showed a loss of DNA methylation in all sequence contexts, but the loss of CpNpG methylation and asymmetric methylation was stronger than that of CpG methylation. Thus the *kyp* methylation phenotype resembles that of *cmt3* more than that of *met1*.

The *kyp* mutants (like *cmt3* mutants) did not develop the late flowering phenotype characteristic of reactivation of the normally methylated and silenced *FWA* locus^{9,21}. In wild-type plants, *FWA* is methylated within two direct repeats of its promoter at CpG sites (88%) and to a lesser extent at CpNpG sites (20%)²¹. This contrasts with *SUP*, which is methylated predominantly at CpNpG sites⁹. To test whether *kyp* mutants affect CpG methylation at *FWA*, we performed a Southern blot analysis of two methylation-sensitive *Cfo*I restriction sites that contain CpG within their recognition sequences (Fig. 3b). The pattern of enzyme digestion was similar in *kyp-1*, *kyp-2* and *clk-st*, showing that neither *kyp* allele affects CpG methylation. We assayed CpNpG methylation at *FWA* with the

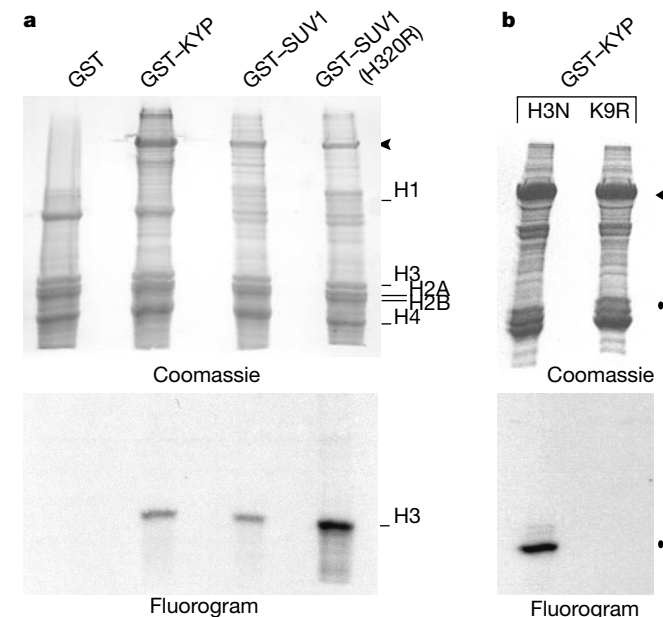


Figure 2 Methyltransferase activity of KRYPTONITE. **a**, Methyltransferase activity of GST, GST–*KYP* (amino acids 291–624), GST–SUV1 (SUV39h1, amino acids 82–412), and GST–SUV1(H320R)¹ fusion proteins with histone substrates and the methyl donor *S*-adenosyl-(methyl-¹⁴C)-L-methionine. Arrowhead marks Coomassie-stained purified proteins (top). Individual histones are indicated. Fluorography (bottom) indicates methyltransferase specificity for histone H3. **b**, GST–*KYP* methyltransferase assays using GST–H3 fusion protein substrates, H3N (57-amino-acid H3N terminus) and K9R (H3N terminus with a mutation to arginine at Lys 9). Arrowhead indicates GST–*KYP* fusion protein. Bullet indicates GST–H3 fusion proteins. Fluorogram shows that GST–*KYP* methylates wild-type but not Lys 9-mutant fusion proteins.

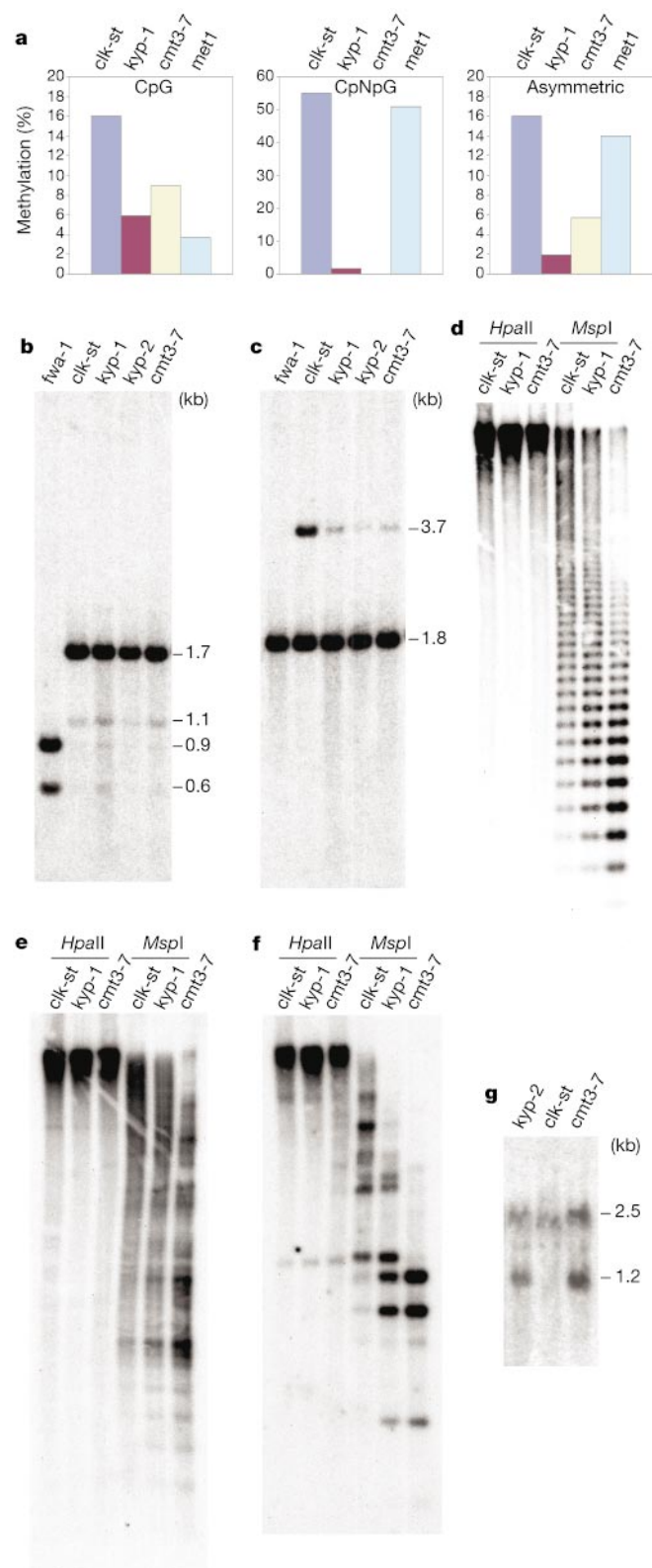


Figure 3 Effect of *kryptonite* on DNA methylation and retrotransposon activation. **a**, Bisulphite sequence analysis showing the methylation of *SUP* in different mutant backgrounds. **b**, **c**, Southern blot of *CfoI*- (**b**) and *BglII*-digested (**c**) genomic DNAs probed with *FWA*²¹. The hypomethylated *fwa-1* mutant²¹ is included as an unmethylated control. **d-f**, Southern blot of *HpaII*- and *MspI*-digested genomic DNAs probed with a 180-base-pair centromeric repeat (**d**)²², an Athila long terminal repeat (**e**)⁹ or Ta3 (**f**)⁹. **g**, RNA blot hybridized with a TSI probe.

methylation-sensitive enzyme *BglII*, and found that *kyp-1*, *kyp-2* and *cmt3-7* all showed a significant increase in digestion relative to the control strain *clk-st*. Thus, like *cmt3* mutants, the *kyp* mutants decreased CpNpG methylation but not CpG methylation of *FWA*. We confirmed these results by bisulphite sequencing the *FWA* locus from *kyp-1* plants. We found levels of CpG methylation comparable to wild type, but lower CpNpG methylation.

We next studied the effect of *kyp* on satellite methylation using a 180-base-pair centromere repeat probe²² and the isoschizomers *HpaII* and *MspI*, which recognize 5'-CCGG-3'. *HpaII* is inhibited by methylation of either the inner (CG) or the outer (CCG) cytosine of the recognition site, whereas *MspI* is sensitive only to methylation of the outer cytosine, and thus detects CpNpG methylation. Similar to *cmt3* mutants⁹, *kyp-1* and *kyp-2* mutants show a small but reproducible increase in *MspI* digestion but not *HpaII* digestion, indicating that *kyp* mutants reduce CpNpG but not CpG methylation at centromeric repeat sequences (Fig. 3d). We also detected increased *MspI* but not *HpaII* cleavage in *kyp-1*, *kyp-2* and *cmt3-7* mutants at two additional sequences tested, a highly repetitive Athila retrotransposon long terminal repeat sequence (Fig. 3e)⁹ and the single-copy Ta3 retrotransposon (Fig. 3f).

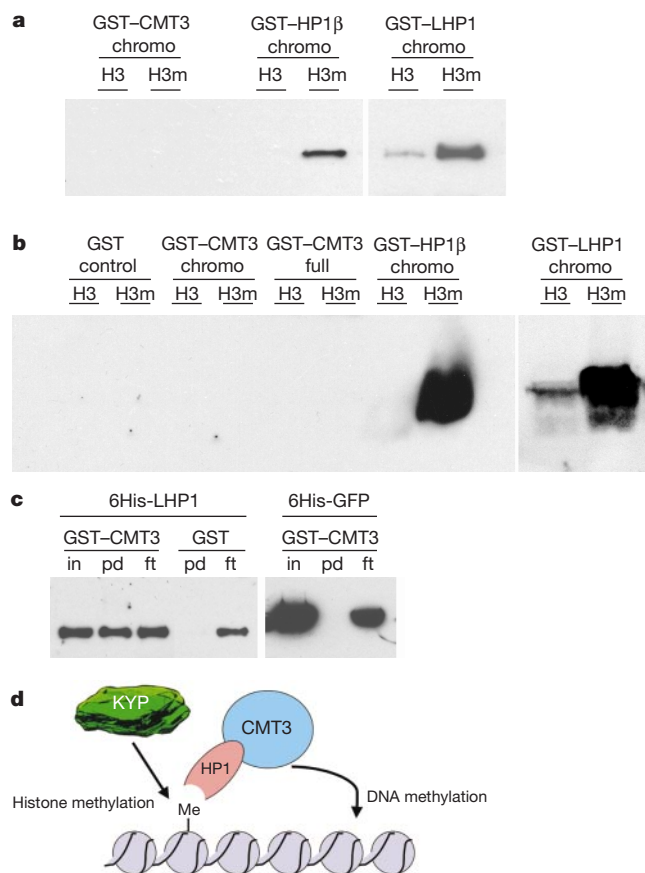


Figure 4 Interaction of CMT3 with histones and LHP1. **a**, Histone H3 peptide pull-down assays. Equivalent amounts of the indicated purified GST fusion proteins were mixed with matrices containing either unmodified histone H3 N-terminal peptides (H3) or Lys 9-methylated peptides (H3m)⁹, and bound proteins were detected with anti-GST antibodies. **b**, Biotinylated H3 N-terminal peptides (H3) or biotinylated Lys 9-methylated peptides (H3m) were mixed with glutathione-agarose-bound GST fusion proteins, and bound proteins were detected with streptavidin horseradish peroxidase (HRP). **c**, GST-CMT3 pull-down assays. Purified His-tagged LHP1 or His-tagged GFP fusion proteins were mixed with glutathione-agarose-bound GST fusion proteins, and bound proteins were detected with INDIA HisProbe-HRP. Approximately 5% of input proteins (in), 100% of bound proteins (pd), and 5% of flow-through (ft) fractions are shown. **d**, Model showing the proposed interaction of CMT3 with methylated histones by binding to LHP1.

Thus the methylation phenotype of *kyp* resembles that of *cmt3*: a loss of CpNpG methylation. However, the effect of *kyp* on CpNpG methylation was weaker than that of *cmt3-7* at most loci tested (Fig. 3). Furthermore, *kyp-1* showed a greater loss of CpG and asymmetric methylation at *SUP* (Fig. 3a), suggesting that, at some loci, its effects on DNA methylation are more general than those of *cmt3-7*.

Similar to the *cmt3* mutants, we found that *kyp* mutants induced a low level of expression of two retrotransposon-related sequences, an Athila sequence called TSI (ref. 23) (Fig. 3g) and the Ta3 sequence⁹ (not shown). These results further demonstrate that *KYP* and *CMT3* affect silencing at an overlapping set of loci.

The close similarity in methylation and silencing phenotypes of *kyp* and *cmt3* mutants prompted us to examine the relationship between *CMT3* and methylated histones. *CMT3*, like *HP1*, contains a chromo domain²⁴, suggesting the possibility that *CMT3* binds directly to methylated H3 Lys 9, thereby directing the methylation of CpNpG sites. To examine this possibility, we first tested for the ability of an *in vitro* translated *CMT3* chromo domain (amino acids 329–490) and full-length *in vitro* translated *CMT3* to bind matrices containing either K9 di-methylated or unmethylated H3 N-terminal peptides. Under previously published conditions⁵, we did not detect specific binding to either matrix. Using a different approach, we expressed full-length *CMT3*, the *CMT3* chromo domain (amino acids 366–448), and a mouse *HP1 β* chromo domain (amino acids 10–80) as GST fusions in *Escherichia coli*. The *HP1 β* chromo domain fusion protein bound specifically to a matrix containing methylated histone H3 peptides, but the full-length *CMT3* and the *CMT3* chromo domain fusion proteins did not (Fig. 4a and not shown). Finally, we reversed the experiment and tested whether methylated or unmethylated biotinylated peptides could bind to the same GST fusion proteins affixed to glutathione-agarose beads. The H3 Lys 9-methylated peptide bound to the *HP1 β* chromo domain fusion but not to the *CMT3* fusion proteins (Fig. 4b). Thus, unlike *HP1*, the *CMT3* chromo domain did not bind specifically to methylated histone H3 under the conditions tested. Several other chromo domains also do not bind methylated histones, including those in the proteins murine PC1 and Suv39h1 (ref. 5), M33 and Mi2 (ref. 6), and Esa1 (ref. 7). Furthermore, chromo domains are proposed to have other functions such as binding to RNA in the case of the histone acetyltransferase MOP²⁵.

Arabidopsis possesses a homologue of *HP1* (*LHP1*) that, like *HP1* from other organisms, localizes to discrete subnuclear foci and homodimerizes through its chromo shadow domain²⁶. We found that the *LHP1* chromo domain (a GST fusion protein containing *LHP1* amino acids 54–183) specifically binds to methylated H3 Lys 9 (Fig. 4a, b). Thus a second possible explanation for the resemblance of *kyp* and *cmt3* is that *CMT3* interacts with *LHP1*, thereby targeting *CMT3* to methylated histones. To test this, we expressed His-tagged full-length *LHP1* and performed pull-down assays with GST–*CMT3*. Figure 4c shows that 6His–*LHP1* bound GST–*CMT3* but not GST alone. A control protein, 6His–GFP, did not bind to the GST–*CMT3* matrix. We also found that *in vitro* translated full-length *CMT3* bound to a GST–*LHP1* fusion protein but not to GST alone, and that *in vitro* translated full-length *LHP1* bound to GST–*CMT3* but not GST alone (not shown). Therefore, *CMT3* binds to *LHP1* *in vitro*. This suggests a model (Fig. 4d) in which *CMT3* is targeted to methylated chromatin through an interaction with *LHP1*. Human *HP1* can distinguish between different types of H3 Lys 9-methylated heterochromatin, binding *in vivo* to constitutive heterochromatin such as centromeric regions, but not to facultative heterochromatin such as the inactive X chromosome²⁷. Therefore, the interaction of *CMT3* with *LHP1* may serve to target CpNpG methylation to particular chromosomal domains.

Our results show that the *KYP* histone H3 Lys 9 methyltransferase is required for maintenance of DNA methylation. Studies of the DIM5 H3 methyltransferase in *Neurospora* demonstrated an intimate relationship between histone methylation and DNA methyl-

ation, because loss of DIM5, as well as mutation of Lys 9 of histone H3, resulted in a complete loss of DNA methylation *in vivo*⁸. Our results extend these findings to the plant kingdom, suggesting that a relationship between histone methylation and DNA methylation is conserved throughout higher eukaryotes. However, unlike the general loss of DNA methylation observed in *dim5SQ14*, *kyp* mutants show a specific loss of CpNpG methylation, a phenotype similar to that of loss-of-function alleles of the *CMT3* DNA methyltransferase gene. One possible explanation for this finding is that, in plants, only CpNpG methylation is controlled by histone H3 Lys 9 methylation. Alternatively, the more subtle effects of *kyp* mutants on DNA methylation might be explained by the presence of eight additional *KYP* homologues in the *Arabidopsis* genome¹⁶, one or more of which might function in the control of DNA methylation in other contexts such as at CpG sites.

Our data suggest that loss of CpNpG methylation in the *kyp* mutants results from lack of targeting of *CMT3* to methylated histones, possibly through an interaction with an *Arabidopsis* *HP1* homologue. The central role that *HP1* plays in chromatin biology of animals and fungi suggests that recruitment of DNA methyltransferases to chromatin by *HP1* could be a general eukaryotic phenomenon.

Methods

Polymerase chain reaction (PCR) primers and PCR-based molecular markers are listed in Supplementary Information.

Characterization of the *kryptonite* alleles

We cloned the *KYP* gene by fine-scale mapping the *kyp-2* allele and sequencing the candidate genes. *kyp-2* was mapped by crossing to a Columbia *sup-2* strain as described⁹. We analysed 1,436 chromosomes (718 *kyp-2* homozygous F₂ plants) with a series of PCR-based molecular markers to narrow the *KYP*-containing region to between a marker on bacterial artificial chromosome (BAC) clone sequence MXE10 and one on BAC clone sequence MUA22. This defined an interval of 141 kilobases (kb) that contained only one additional BAC clone sequence, MAC12, which contains the *KYP* gene (MAC12.7; GenBank accession number AB005230). A description of the molecular nature of each *kyp* allele is provided in Supplementary Information.

Histone methyltransferase assays

A GST–*KYP* fusion construct was made by cloning *KYP* (amino acids 281–625) into the *Bam*HI and *Eco*R1 sites of pGEX2TK by RT–PCR (PCR with reverse transcription of RNA) from Dnase-treated total RNA of Landsberg erecta (Ler)-strain inflorescences, and confirmed by sequencing. The GST–SUV1 (82–412) and GST–SUV1(H320R) (a mutant version in which histidine 320 is mutated to arginine) constructs¹ were a gift of T. Jenuwein. Recombinant proteins were expressed and purified on glutathione-agarose beads (Pierce) as described¹. *In vitro* histone methyltransferase reactions were performed as described¹ and proteins were separated by 18% SDS–PAGE and visualized by Coomassie staining and fluorography. Histone methyltransferase reactions on GST–H3 fusion proteins¹⁷ (a gift of Y. Shinkai) were performed using similar reaction conditions.

Analysis of DNA methylation

Genomic DNA was isolated from leaves of 4-week-old flowering plants of all genotypes. For the results reported in Fig. 3a, bisulphite sequencing on the top strand of a 1,028-nucleotide region of *SUP* was performed as previously described⁹, with 15 independent cloned PCR products of each region analysed. *kyp-1* and *cmt3-7* were in a homozygous *clk-st* background and *met1* was a previously described line in a *clk* background⁹. Detailed data supporting the graphical presentation in Fig. 3a can be found in Supplementary Information Table 1. Bisulphite sequencing of *FWA* was as previously described²¹, and the data reported in the text are from an analysis of eight cloned PCR products.

RNA blot analysis

For Fig. 3g, total RNA was isolated from 4-week-old plants of each genotype, 40 μ g of RNA was loaded per lane, and a blot was probed with a TSI probe⁹. As we observed previously with this probe⁹, there was a less-abundant TSI transcript of about 2.4 kb detected in the *clk-st* strain.

Histone peptide binding assays

GST fusion constructs were made by cloning fragments of *CMT3*, *HP1 β* and *LHP1* into the *Bam*HI and *Eco*R1 sites of pGEX2TK, and clones were confirmed by DNA sequencing. Affinity matrices were generated by coupling unmodified H3 N-terminal (residues 1–20) or K9 dimethyl peptides (a gift of T. Jenuwein) to SulfoLink Coupling Gel (Pierce) as described by the manufacturer. Binding conditions, buffers, and wash conditions for

experiments shown in Fig. 4a, b were exactly as described⁵, except that the binding and washing buffer for the GST–LHP1 results shown in Fig. 4a consisted of 50 mM sodium phosphate and 25 mM NaCl at pH 6.0. For Fig. 4a, bound proteins were separated by 4–20% gradient SDS–PAGE, blotted, and detected with an anti-GST monoclonal antibody (Pierce). For Fig. 4b, the bound biotinylated peptides (Upstate Biotechnology) were separated by 18% SDS–PAGE, blotted, and detected with streptavidin HRP conjugate (Upstate Biotechnology).

Interaction of CMT3 with LHP1

A six-histidine fusion construct was made by cloning full-length LHP1 into the *Xho*I and *Pst*I sites of pRSETB, and expressed in *E. coli* strain BL21. Proteins were purified on Ni-NTA-agarose (Qiagen) and eluted with 100 mM imidazole before mixing with glutathione-agarose-bound GST fusion proteins. Binding and wash buffers were the same as in the peptide binding assays⁵. Bound proteins were separated by 4–20% gradient SDS–PAGE, blotted, and detected with INDIA HisProbe-HRP (Pierce).

Received 15 January; accepted 4 March 2002.

Published online 17 March 2002, DOI 10.1038/nature731.

1. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
2. Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110–113 (2001).
3. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).
4. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074–1080 (2001).
5. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
6. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
7. Jacobs, S. A. *et al.* Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J.* **20**, 5232–5241 (2001).
8. Tamaru, H. & Selker, E. U. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414**, 277–283 (2001).
9. Lindroth, A. M. *et al.* Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. *Science* **292**, 2077–2080 (2001).
10. Tschiersch, B. *et al.* The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* **13**, 3822–3831 (1994).
11. Allshire, R. C., Nimmo, E. R., Ekwil, K., Javerzat, J. P. & Cranston, G. Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev.* **9**, 218–233 (1995).
12. Ivanova, A. V., Bonaduce, M. J., Ivanov, S. V. & Klar, A. J. The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. *Nature Genet.* **19**, 192–195 (1998).
13. Peters, A. H. *et al.* Loss of the *suv39h* histone methyltransferase impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).
14. Jacobsen, S. E. & Meyerowitz, E. M. Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. *Science* **277**, 1100–1103 (1997).
15. Bartee, L., Malagnac, F. & Bender, J. *Arabidopsis* cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. *Genes Dev.* **15**, 1753–1758 (2001).
16. Baumbusch, L. O. *et al.* The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res.* **29**, 4319–4333 (2001).
17. Tachibana, M., Sugimoto, K., Fukushima, T. & Shinkai, Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* **276**, 25309–25317 (2001).
18. Finnegan, E. J., Peacock, W. J. & Dennis, E. S. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. *Proc. Natl Acad. Sci. USA* **93**, 8449–8454 (1996).
19. Ronemus, M. J., Galbiati, M., Ticknor, C., Chen, J. & Dellaporta, S. L. Demethylation-induced developmental pleiotropy in *Arabidopsis*. *Science* **273**, 654–657 (1996).
20. Kishimoto, N. *et al.* Site specificity of the *Arabidopsis* MET1 DNA methyltransferase demonstrated through hypermethylation of the superman locus. *Plant Mol. Biol.* **46**, 171–183 (2001).
21. Soppe, W. J. *et al.* The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. *Mol. Cell* **6**, 791–802 (2000).
22. Vongs, A., Kakutani, T., Martienssen, R. A. & Richards, E. J. *Arabidopsis thaliana* DNA methylation mutants. *Science* **260**, 1926–1928 (1993).
23. Steimer, A. Endogenous targets of transcriptional gene silencing in *Arabidopsis*. *Plant Cell* **12**, 1165–1178 (2000).
24. Henikoff, S. & Comai, L. A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in *Arabidopsis*. *Genetics* **149**, 307–318 (1998).
25. Akhtar, A., Zink, D. & Becker, P. B. Chromodomains are protein–RNA interaction modules. *Nature* **407**, 405–409 (2000).
26. Gaudin, V. *et al.* Mutations in LIKE HETEROCHROMATIN PROTEIN 1 affect flowering time and plant architecture in *Arabidopsis*. *Development* **128**, 4847–4858 (2001).
27. Peters, A. H. F. M. *et al.* Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nature Genet.* **30**, 77–80 (2002).

Supplementary Information accompanies the paper on Nature's website (<http://www.nature.com>).

Acknowledgements

We thank T. Jenuwein for the GST–Suv constructs and H3 N-terminal peptides, Y. Shinkai for the H3 N-terminal GST fusion constructs, A. Kouzarides for an HP1 construct, and S. Peyvandi for technical assistance. This work was supported by grants from the National Institutes of Health, the Beckman Young Investigator programme, and the Searle Scholars Foundation to S.E.J. J.P.J. was supported by an NIH training grant and A.M.L. by a post-doctoral fellowship from the Damon Runyon Walter Winchel Foundation.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to S.E.J. (e-mail: jacobsen@ucla.edu).

p63 and p73 are required for p53-dependent apoptosis in response to DNA damage

Elsa R. Flores*, Kenneth Y. Tsai*†, Denise Crowley*§, Shomit Sengupta*§, Annie Yang‡, Frank McKeon‡ & Tyler Jacks*§

*Massachusetts Institute of Technology, Department of Biology and Center for Cancer Research, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

†Harvard–Massachusetts Institute of Technology Division of Health Sciences and Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

‡Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA

§Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA

The tumour-suppressor gene *p53* is frequently mutated in human cancers and is important in the cellular response to DNA damage^{1,2}. Although the *p53* family members *p63* and *p73* are structurally related to *p53*, they have not been directly linked to tumour suppression, although they have been implicated in apoptosis^{3–9}. Given the similarity between this family of genes and the ability of *p63* and *p73* to transactivate *p53* target genes^{10,11}, we explore here their role in DNA damage-induced apoptosis. Mouse embryo fibroblasts deficient for one or a combination of *p53* family members were sensitized to undergo apoptosis through the expression of the adenovirus E1A onco-gene^{12–14}. While using the E1A system facilitated our ability to perform biochemical analyses, we also examined the functions of *p63* and *p73* using an *in vivo* system in which apoptosis has been shown to be dependent on *p53*. Using both systems, we show here that the combined loss of *p63* and *p73* results in the failure of cells containing functional *p53* to undergo apoptosis in response to DNA damage.

Previous work has shown that *p53*-deficient E1A mouse embryo fibroblasts (MEFs) treated with DNA-damaging agents are highly resistant to apoptosis^{12–14}. To determine whether *p63* and/or *p73* are involved in apoptosis induced by DNA damage, *p63*-deficient (*p63*^{−/−}) and *p73*-deficient (*p73*^{−/−}) E1A-expressing MEFs were generated and treated with doxorubicin for 0, 6, 12, 24 and 48 h (Fig. 1A; Supplementary Information, Fig. S1a), stained with annexin V coupled to fluorescein isothiocyanate, and analysed by flow cytometry. E1A MEFs lacking *p63* or *p73* exhibited a partial resistance to apoptosis in response to DNA damage; 70% of the *p63*^{−/−} E1A MEFs and 80% of the *p73*^{−/−} E1A MEFs were viable, compared with 42% of the wild-type E1A MEFs at 12 h. At 24 h, 50% of the *p63*^{−/−} E1A MEFs and 65% of the *p73*^{−/−} E1A MEFs were viable, compared with 5% of the wild-type E1A MEFs (Fig. 1A).

Because *p63*-deficient and *p73*-deficient E1A MEFs exhibited a partial resistance to apoptosis, we hypothesized that they might cooperate with *p53* or with each other in the DNA damage response. Therefore, double-homozygous E1A MEFs deficient for all pairs of combinations (*p53*^{−/−}; *p63*^{−/−}, *p53*^{−/−}; *p73*^{−/−}, *p63*^{−/−}; *p73*^{−/−}) were generated and treated with doxorubicin. As shown in Fig. 1A, all the double-knockout cells, including the *p63*^{−/−}; *p73*^{−/−} E1A MEFs, were resistant to apoptosis in response to DNA damage. The *p53*^{−/−}; *p63*^{−/−} and *p53*^{−/−}; *p73*^{−/−} E1A MEFs were more resistant

than the $p53^{-/-}$ E1A MEFs, whereas the $p63^{-/-}$; $p73^{-/-}$ E1A MEFs were as resistant as $p53^{-/-}$ E1A MEFs (Fig. 1A). Cisplatin and γ -irradiation were also tested on E1A MEFs of the different genotypes with similar results (Supplementary Information, Fig. S1). These findings indicate that even in the presence of $p53$, the E1A MEFs deficient for both $p63$ and $p73$ are resistant to apoptosis, suggesting that these genes might act together with $p53$ or in an obligatory parallel pathway to induce apoptosis subsequent to DNA damage.

Several isoforms have been identified for $p63$ and $p73$, including some containing (TAp63 and TAp73) and lacking (Δ Np63 and Δ Np73) an amino-terminal transactivation domain⁴. The $p63^{-/-}$; $p73^{-/-}$ E1A MEFs carry mutations in the central DNA-binding domains and are therefore deficient for all of these isoforms. To determine whether the transactivation domain of these proteins is important in the response to DNA damage, wild-type, $p53^{-/-}$ and $p63^{-/-}$; $p73^{-/-}$ E1A MEFs were transfected with expression vectors for TAp63 α , TAp73 α , Δ Np63 α or Δ Np73 α . After 24 h the cells were treated with 0.34 μ M doxorubicin for 12 h, and apoptosis was detected by staining with annexin V. Protein expression of the transfected plasmids was detected by western blotting (Supplementary Information, Fig. S1d). Wild-type E1A MEFs transfected with any of the expression vectors underwent apoptosis in response to doxorubicin, whereas $p53^{-/-}$ E1A MEFs transfected with these constructs and treated with doxorubicin did not. Transfection of TAp63 α or TAp73 α into the $p63^{-/-}$; $p73^{-/-}$ E1A MEFs caused a small but consistent increase in apoptosis after treatment with doxorubicin, whereas transfection with Δ Np63 α or Δ Np73 α had no effect in this system (Supplementary Information, Fig. S1e). These data suggest that the transactivation domains of $p63$ and $p73$

are important for DNA-damage induced apoptosis.

$p53$ has been shown to be required for irradiation-induced apoptosis in the developing nervous system^{15–17}. To examine the requirements for $p63$ and $p73$ in this setting, control and $p63^{-/-}$; $p73^{-/-}$ day-13.5 embryos were analysed. Wild-type, $p53^{-/-}$, $p63^{-/-}$, $p73^{-/-}$ and $p63^{-/-}$; $p73^{-/-}$ embryos were treated with γ -irradiation (5 Gy) *in utero*, harvested 5 h later and analysed for apoptosis by end labelling *in situ*. Figure 1B, C shows that there was a striking difference between the level of apoptosis detected in the γ -irradiated central nervous system (CNS) of the wild-type embryos ($29 \pm 8\%$) compared with the $p53^{-/-}$ ($0.18 \pm 0.1\%$) and the $p63^{-/-}$; $p73^{-/-}$ embryos ($0.2 \pm 0.1\%$) (Fig. 1Bb, Bd, Bf, C). The level of apoptosis in the $p63^{-/-}$; $p73^{-/-}$ embryos was similar to that found in $p53^{-/-}$ embryos, indicating that $p63$ and $p73$ are required for $p53$ -dependent apoptosis in this system as well. Both the $p63^{-/-}$ and $p73^{-/-}$ embryos displayed an intermediate phenotype: $7.5 \pm 3.0\%$ and $8.0 \pm 3.0\%$ apoptotic cells, respectively (Fig. 1Bg, Bh, C). No apoptosis was detected in untreated wild-type, $p53^{-/-}$, and $p63^{-/-}$, $p73^{-/-}$ embryos (Fig. 1Ba, Bc, Be). The number of apoptotic cells was divided by the number of total cells to derive the percentage of apoptotic cells in each sample. These results indicate that both $p63$ and $p73$ are required for $p53$ -dependent apoptosis *in vivo* and support the results previously derived from E1A expressing MEFs.

Given that the $p63^{-/-}$; $p73^{-/-}$ MEFs expressing E1A were as resistant to apoptosis as cells lacking $p53$ itself, experiments were performed to determine the function of $p53$ in the double-mutant cells. In all cells containing functional $p53$, the $p53$ protein was expressed at similar levels and was induced on treatment with

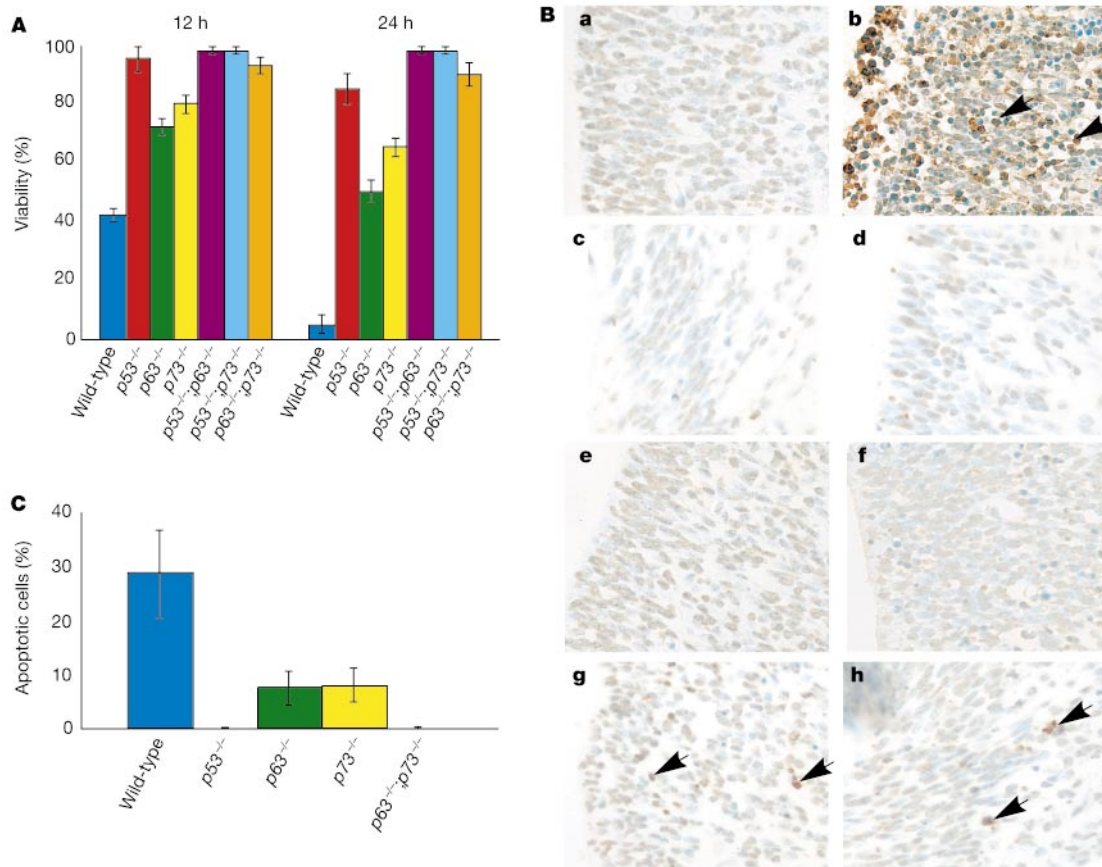


Figure 1 E1A MEFs and the CNS from 13.5-day embryos mutant for $p53$ -family proteins are resistant to apoptosis in response to DNA damage. **A**, Histogram representing the percentage of viable cells scored by annexin 12 and 24 h after treatment with 0.34 μ M doxorubicin. **B**, CNS of embryos assayed for apoptosis by using FRAGEL. **a, c, e**, Untreated embryos: wild type (**a**), $p53^{-/-}$ (**c**) and $p63^{-/-}$; $p73^{-/-}$ (**e**).

b, d, f, g, h Embryos treated *in utero* with γ -irradiation (5 Gy) for 5 h: wild type (**b**), $p53^{-/-}$ (**d**), $p63^{-/-}$; $p73^{-/-}$ (**f**), $p63^{-/-}$ (**g**) and $p73^{-/-}$ (**h**). Positive cells are indicated by arrows. $p63^{-/-}$; $p73^{-/-}$ embryos were histologically indistinguishable from wild-type embryos (see description in Supplementary Information). **C**, Histogram showing the percentages of apoptotic cells after γ -irradiation for 5 h from the analysis in **B**.

doxorubicin for 12 h, as shown by western blot analysis (Fig. 2a). We also determined that p63 is induced on treatment with doxorubicin. By western blot analysis, wild-type, $p53^{-/-}$ and $p53^{-/-}; p73^{-/-}$ E1A-expressing cells exhibited a striking increase in the amount of p63 after treatment with doxorubicin for 12 h (Fig. 2b). Antibodies against mouse p73 were not of sufficient quality to determine p73 induction, although endogenous human p73 has been shown to be induced by E1A and cisplatin^{6,7,9,7,18}. Therefore, p63 and p73 can be induced after forms of stress, including DNA damage and oncogenic stimuli.

Upon DNA damage, p53 is stabilized and accumulates in the nucleus. Indirect immunofluorescence for p53 was performed on wild-type and $p63^{-/-}; p73^{-/-}$ E1A MEFs after treatment with doxorubicin, cisplatin or γ -irradiation. The extent and pattern of p53 nuclear staining in the $p63^{-/-}; p73^{-/-}$ E1A MEFs were comparable to those of wild-type E1A cells (data not shown). Additionally, the localization of p63 was tested in wild-type, $p53^{-/-}$ and $p63^{-/-}; p73^{-/-}$ E1A MEFs after treatment with doxorubicin for 6 h (Supplementary Information, Fig. S2). Similarly to p53, p63 nuclear staining was increased under these conditions, consistent

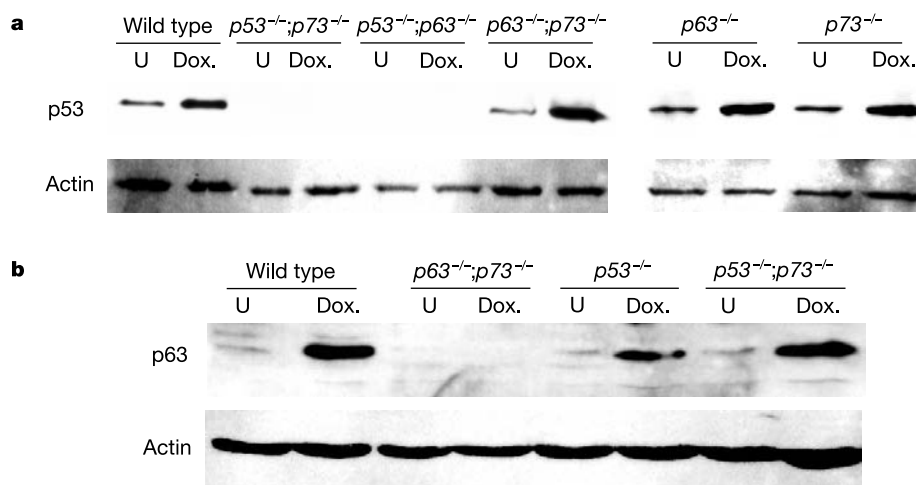


Figure 2 p53 and p63 are induced in E1A MEFs after treatment with DNA-damaging agents. Western blots are shown containing whole cell extracts from MEFs expressing E1A. **a**, p53 is induced in wild-type, $p63^{-/-}$; $p73^{-/-}$, $p63^{-/-}$ and $p73^{-/-}$ E1A MEFs.

b, p63 is induced in wild-type, $p53^{-/-}$ and $p53^{-/-}; p73^{-/-}$ E1A MEFs 12 h after treatment with 0.34 μ M doxorubicin (Dox.). Actin was used as a loading control. U, untreated.

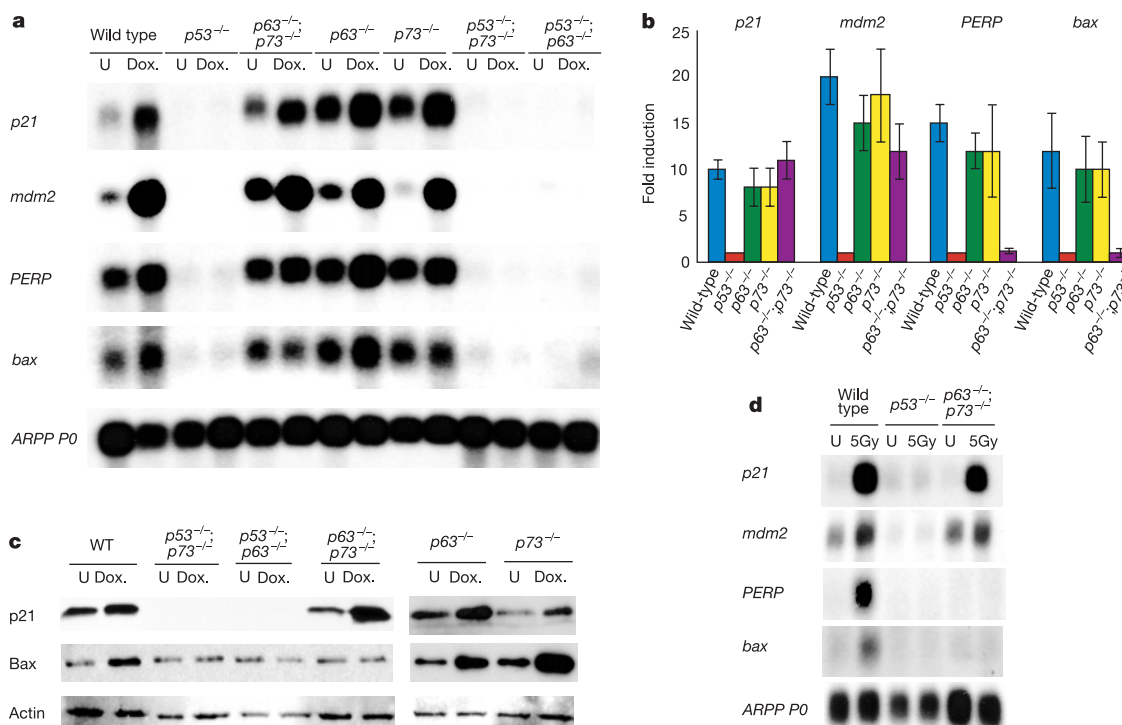


Figure 3 p53 target genes are induced differently in $p63^{-/-}; p73^{-/-}$ E1A MEFs and the brain of 13.5-day embryos. **a**, Northern blot analysis of total RNA extracted from E1A MEFs treated with 0.34 μ M doxorubicin (Dox.) for 12 h and probed with *mdm2*, *p21*, *PERP*, *bax* or *ARPP P0*. **b**, Histogram indicating fold induction of the p53 target genes

analysed in **a** calculated with the use of ARPP P0 and the PhosphorImager. **c**, Western blot analysis of E1A MEFs treated for 12 h with 0.34 μ M doxorubicin and probed with anti-p21 and anti-Bax. **d**, Northern blot analysis of total RNA extracted from brains of 13.5-day embryos 5 h after treatment with γ -irradiation (5 Gy). U, untreated.

with its having a function in the transcriptional activation of target genes involved in the DNA damage response.

For the further characterization of MEFs expressing E1A and deficient for either *p63* or *p73* or both, the expression of known *p53* target genes was examined. Specifically, the *p21*, *mdm2*, *bax* and *PERP* genes were analysed by northern blotting both before and 12 h after treatment with 0.34 μ M doxorubicin^{19–21}. Figure 3a, b shows that the induction of *p21* and *mdm2* was comparable in wild-type, *p63*^{−/−}, *p73*^{−/−} and *p63*^{−/−}; *p73*^{−/−} E1A MEFs. Interestingly, whereas *bax* and *PERP* were also induced to similar extents in wild-type, *p63*^{−/−} and *p73*^{−/−} E1A MEFs, these genes were not induced in the *p63*^{−/−}; *p73*^{−/−} E1A MEFs (see Fig. 3b). Western blot analysis was performed for *p21* and *bax* to confirm these results at the level of protein expression (Fig. 3c). These data show that some *p53* targets are regulated normally in *p63*^{−/−}; *p73*^{−/−} E1A MEFs, whereas genes linked to the apoptotic response were specifically affected.

To determine whether *p53* target genes were differentially expressed in the CNS of developing *p63*^{−/−}; *p73*^{−/−} day-13.5 embryos, the brain was microdissected from control wild-type, *p53*^{−/−} and *p63*^{−/−}; *p73*^{−/−} embryos or embryos treated *in utero* with γ -irradiation (5 Gy). At 5 h after treatment, total RNA was isolated from each brain for northern blot analysis. Again, *p21*, *mdm2*, *PERP* and *bax* were found to be induced in wild-type embryos, whereas only *p21* and *mdm2* were induced in *p63*^{−/−}; *p73*^{−/−} embryos, confirming the finding in the E1A MEFs (Fig. 3d).

Given that *p53* and *p63* are present, induced and enriched in the nuclei of E1A MEFs after DNA damage, we assayed for *p53* and *p63* occupancy of specific promoter elements *in vivo* by chromatin immunoprecipitation (ChIP) analysis for *p53* target genes. We

examined the *mdm2* and *p21* promoters, which had been shown to be induced in both wild-type and *p63*^{−/−}; *p73*^{−/−} E1A MEFs by northern analysis, and *bax* and *PERP*, which were induced in wild-type but not *p63*^{−/−}; *p73*^{−/−} E1A MEFs (Fig. 3a–c). Additionally, the *NOXA* promoter was used as another example of a *p53* target gene believed to be important for *p53*-dependent apoptosis²². Figure 4a, b shows that after DNA damage there was increased association of *p53* and *p63* with the *p21*, *mdm2*, *bax*, *PERP* and *NOXA* promoters in wild-type MEFs expressing E1A. Similarly, increased association of *p53* with the *p21* and *mdm2* promoters was evident in *p63*^{−/−}; *p73*^{−/−} E1A MEFs under these conditions, although at somewhat lower levels than in wild-type E1A MEFs. In striking contrast, we observed no association of *p53* with the *PERP*, *NOXA* and *bax* promoters after DNA damage in *p63*^{−/−}; *p73*^{−/−} E1A cells. These results correlate with the differential induction of *p53* target genes obtained in the northern blot analyses (Fig. 3a–d), and indicate that after DNA damage *p63* and *p73* might regulate the ability of *p53* to bind at certain selected promoters. Interestingly, *p63* was also found at the *bax*, *PERP* and *NOXA* promoters in *p53*^{−/−} E1A MEFs after treatment with doxorubicin for 12 h (Fig. 4b), whereas *p63* was not apparent at the *mdm2* and *p21* promoters in these cells. Notably, *p63* was found to bind both the 5' and 3' *p53*-binding sites on the *PERP* promoter²¹ and was more strongly associated with the 5' site, whereas *p53* was detected only at the 3' site, suggesting a complex regulation of this gene by the *p53* family members. These results indicate that *p63* is present at *p53* target genes after DNA damage and, importantly, that it is present in the absence of *p53* at target genes involved in the apoptotic process. (For a complete set of controls, see Supplementary Information, Fig. S4.)

p53 has long been recognized as central to the induction of apoptosis in response to DNA damage, a function thought to be critical for tumour suppression and the response of tumour cells to chemotherapeutic agents. Although our data do not diminish the significance of *p53* in this process, they do establish a role for *p63* and *p73*. The combined absence of *p63* and *p73* severely impaired the induction of *p53*-dependent apoptosis in response to DNA damage in E1A-expressing cells and in the developing mouse CNS. This can be explained by the inability of *p53* to bind the promoters of apoptosis-associated target genes and to upregulate their transcription in *p63*^{−/−}; *p73*^{−/−} E1A cells and the developing CNS. There are two pieces of evidence that *p63* and *p73* are important players in recruiting *p53* to promoters at apoptosis-related genes. First, *p63* is present at such promoters, even in the absence of *p53*. Second, *p53* is not present at these promoters in the absence of *p63* and *p73*. Our data support the notion that there might be two classes of *p53*-family target genes. One class includes genes such as *mdm2* and *p21*, which *p53* regulates in the presence or absence of *p63* and *p73*. The other includes genes such as *PERP*, *bax* and *NOXA*, for which genetic and biochemical data indicate a requirement for *p63* or *p73* for *p53* to be recruited and function properly. These data show that *p63* and *p73* are important components of the response to DNA damage, and may portend a greater role for these proteins in tumour suppression and chemosensitivity. □

Methods

Cell culture

MEFs were derived from 13.5-day embryos from the following mouse crosses: *p53*^{−/−} embryos were derived from *p53*^{+/−} crosses maintained on a 129/SvJae background; *p63*^{−/−} and *p73*^{−/−} embryos were similarly derived from *p63*^{+/−} (maintained on a mixed C57BL/6 × 129/SvJae) and *p73*^{+/−} crosses (maintained on a 129/SvJae background), respectively. Double-mutant MEFs (*p53*^{−/−}; *p73*^{−/−}, *p53*^{−/−}; *p63*^{−/−} and *p63*^{−/−}; *p73*^{−/−}) were derived by first generating double-heterozygous mice and crossing to generate double-homozygous embryos. The genotypes of the MEFs were determined by polymerase chain reaction (PCR) as described previously^{23–25}. E1A-expressing MEFs of each of the above genotypes were generated by infection with high-titre ecotropic retroviruses derived by using the Phoenix retrovirus packaging system²⁶. At 48 h after infection, E1A-expressing MEFs were selected by the addition of 2 μ g ml^{−1} puromycin for 2 d. For transfection experiments, E1A MEFs were transfected with 0.5 μ g mouse TAp63 α , Δ Np63 α , Δ Np73 α ,

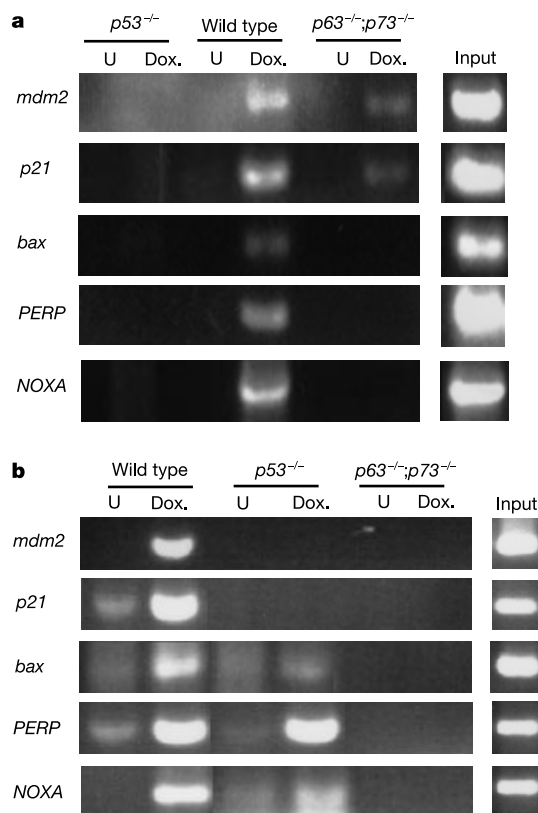


Figure 4 Differential binding of *p53* and *p63* *in vivo* at various promoters in E1A MEFs. **a**, ChIP analysis for the presence of *p53* at the *p21*, *mdm2*, *PERP* (3' site), *NOXA* and *bax* promoters in E1A MEFs before (U) and after treatment for 12 h with 0.34 μ M doxorubicin (Dox.). **b**, ChIP analysis for the presence of *p63* at the *p21*, *mdm2*, *PERP* (5' site), *NOXA* and *bax* promoters in E1A MEFs. Input, total input chromatin from untreated wild-type E1A MEFs. For additional controls see Supplementary Information, Fig. S4.

human Tap73 α or pcDNA3 with the use of FuGene (Roche). At 24 h after transfection, cells were treated with doxorubicin for 12 h and assayed for apoptosis as described below.

Annexin assay

At 24 h after plating, E1A MEFs were treated with 0.34 μ M doxorubicin. Floating and adherent cells were collected 0, 6, 12, 24 and 48 h after treatment, stained with annexin V coupled to fluorescein isothiocyanate (Becton Dickinson) and propidium iodide (50 μ g ml⁻¹) in accordance with the manufacturer's instructions, and analysed with a Becton Dickinson FACScan with the use of CellQuest software. The percentage of viable cells was calculated by scoring for cells that were not annexin-positive (early apoptotic) or not positive for both annexin and propidium iodide (late apoptotic and necrotic).

End labelling *in situ* on 13.5-day embryonic CNS

Wild-type, p53^{-/-}, p63^{-/-}, p73^{-/-} and p53^{-/-}; p73^{-/-} day-13.5 embryos were derived from the same mouse crosses used to derive the MEFs. Pregnant females at 13 d post coitus (p.c.) were treated with γ -irradiation (5 Gy) by using an irradiator with a ¹³⁷Cs source. At 5 h after treatment, embryos were harvested and fixed in 10% neutral buffered formalin overnight. Embryos were processed, embedded in paraffin and sectioned onto slides. End labelling *in situ* was performed on slides with the use of the FRAGEL kit from Oncogene Research Products, in accordance with the manufacturer's protocol. Brown nuclei were scored as positive. To quantify the numbers of apoptotic cells in the CNS of the embryos, brown nuclei in the subventricular zone were counted and divided by the total number of cells in that zone. Four or five embryos per genotype were used in this analysis.

Northern blot analysis

At 24 h after plating, E1A MEFs of each genotype were treated for 12 h with 0.34 μ M doxorubicin. For embryonic brain analysis, embryos were treated *in utero* with γ -irradiation (5 Gy) by using an irradiator with a ¹³⁷Cs source at 13 d post coitus; 5 h after treatment, brains were harvested from the embryos by microdissection. Total RNA was isolated from E1A MEFs or from embryonic brains by using Tri reagent (Sigma) and subjected to electrophoresis on a 1% agarose gel in 1 $\times$ MOPS containing 6% formaldehyde. Blots probed with *mdm2*, *p21* and ARPP P0 were prehybridized and hybridized in ExpressHyb solution at 65 °C (Clontech). Blots probed with *bax* and *PERP* were prehybridized in 0.5 M NaPO₄ pH 7.2, 1 mM EDTA, 1% BSA, 7% SDS and hybridized in 30% formamide, 0.2 M NaPO₄ pH 7.2, 1 mM EDTA, 1% BSA, 7% SDS at 42 °C. Washes were performed at 60 °C in 2 $\times$ SSC, 0.1% SDS for 15 min followed by 0.2 $\times$ SSC, 0.1% SDS for 15 min. Radiolabelled probes from complementary DNAs corresponding to *mdm2*, *p21*, *bax*, *PERP* and ARPP P0 were prepared by random primer labelling with the Prime-It II kit (Stratagene) and [α -³²P]dATP.

Western blot analysis

E1A MEFs treated for 12 h with 0.34 μ M doxorubicin were lysed on ice in lysis buffer (100 mM Tris, 100 mM NaCl, 1% Nonidet P40, protease inhibitor cocktail (Roche)), 100 μ g of protein was loaded on a SDS/10% polyacrylamide gel and transferred to a poly(vinylidene difluoride) membrane (Millipore). Blots were blocked in TBST (10 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% Tween) containing 5% milk and probed with the following primary antibodies: anti-p21 (clone F5; Santa Cruz), anti-p53 (Ab-3; Oncogene Research Products), anti-p63 (clone 4A4; F. McKeon), anti-Bax (clone N-20; Santa Cruz), anti-c-Myc (clone 9E-10; Santa Cruz) and anti-haemagglutinin (clone Y11, Santa Cruz). After incubation with the appropriate secondary antibody conjugated to horseradish peroxidase (Jackson Immunochemicals), detection was performed with the ECL Plus kit (Amersham). To ensure even loading of protein samples, each blot was stripped with strip buffer (125 mM Tris-HCl pH 6.8, 2% SDS, 0.68% 2-mercaptoethanol) at 50 °C for 30 min followed by a 30-min wash in TBST at 25 °C and probed with actin (Santa Cruz).

p63 immunofluorescence

At 24 h after plating, E1A MEFs were treated for 6 h with 0.34 μ M doxorubicin and fixed on coverslips in 1:1 methanol/acetone for 1 min. Cells were rehydrated in 1 $\times$ PBS, blocked in 1% normal goat serum and incubated with a mouse monoclonal anti-p63 antibody (clone 4A4, F. McKeon) at 25 °C for 2 h. Detection was performed with an anti-mouse Texas Red secondary antibody (Molecular Probes) at 25 °C for 2 h.

ChIP assay

ChIP analysis was performed as described previously²⁷. In brief, adherent E1A-expressing MEFs untreated or treated with doxorubicin (0.34 μ M) for 12 h were removed from the dish with trypsin, which was subsequently inactivated with DMEM containing fetal bovine serum. Cellular proteins were crosslinked to chromatin with 1% formaldehyde for 10 min at 25 °C. p53–DNA complexes or p63–DNA complexes were immunoprecipitated from nuclear extracts by using antibodies against mouse p53 (Ab-1 and Ab-4; Oncogene Research Products) or p63 (clone 4A4; F. McKeon), recovered with *Staphylococcus A* cells and treated with proteinase K to purify DNA. PCR was performed by amplifying ~250-base-pair fragments of the promoter sequence.

The following primers were used. *mdm2*: 5' end, 5'-GGTGCCTGGTCCCGGACTCG CCGGG-3'; 3' end, 5'-CCGAGAGGTGCCCCAGGGGTGTCC-3', site located in intron 1; annealing temperature 75 °C. *p21*: 5' end, 5'-CCTTTCTATCAGCCCCAGAGGATACC-3'; 3' end, 5'-GGGACGTCTTAATTATCTGG GGTC-3', at site -2853; annealing temperature 58 °C. *bax*: 5' end, 5'-GATGTTGTAGCC ACCGCGTACAGCC-3'; 3' end, 5'-TTCATGGTAGAGACACTAAGGAGG-3', encompassing sites -518 and -324; annealing temperature 65 °C. *PERP* 5' site: 5' end, 5'-CACACAATCAGAAGGCCTTGG AGGG-3'; 3' end, 5'-TAATACTCCGAGTTGAGGC AGGAAG-3', at site -2097;

annealing temperature 59 °C. *PERP* 3' site: 5' end, 5'-AGCCACTGGGACCTGCTGGTCA CCC-3'; 3' end, 5'-CTACCTGCTGGGTACACCCGG GCGGC-3', at site -218; annealing temperature 68 °C. *NOXA*: 5' end, 5'-CCTGGCCC CACCCACCCACCC-3'; 3' end, 5'-TCAGGGCTATTTTACGCTCTCGGCC-3', at site -174; annealing temperature 65 °C.

Received 9 November 2001; accepted 25 February 2002.

- Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323–331 (1997).
- Evan, G. I. & Vousden, K. H. Proliferation, cell cycle and apoptosis in cancer. *Nature* **411**, 342–348 (2001).
- Yang, A. *et al.* p63, a p53 homolog at 3q27–29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol. Cell* **2**, 305–316 (1998).
- Irwin, M. S. & Kaelin, W. G. Jr Role of the newer p53 family proteins in malignancy. *Apoptosis* **6**, 17–29 (2001).
- Stiewe, T. & Putzer, B. M. Role of the p53-homologue p73 in E2F1-induced apoptosis. *Nature Genet.* **26**, 464–469 (2000).
- Agami, R., Blandino, G., Oren, M. & Shaul, Y. Interaction of c-Abl and p73 α and their collaboration to induce apoptosis. *Nature* **399**, 809–813 (1999).
- Gong, J. G. *et al.* The tyrosine kinase c-Abl regulates p73 in apoptotic response to cisplatin-induced DNA damage. *Nature* **399**, 806–809 (1999).
- Jost, C. A., Marin, M. C. & Kaelin, W. G. Jr p73 is a simian [correction of human] p53-related protein that can induce apoptosis. *Nature* **389**, 191–194 (1997).
- Yuan, Z. M. *et al.* p73 is regulated by tyrosine kinase c-Abl in the apoptotic response to DNA damage. *Nature* **399**, 814–817 (1999).
- Di Como, C. J., Gaiddon, C. & Prives, C. p73 function is inhibited by tumor-derived p53 mutants in mammalian cells. *Mol. Cell. Biol.* **19**, 1438–1449 (1999).
- Gaiddon, C., Lokshin, M., Ahn, J., Zhang, T. & Prives, C. A subset of tumor-derived mutant forms of p53 down-regulate p63 and p73 through a direct interaction with the p53 core domain. *Mol. Cell. Biol.* **21**, 1874–1887 (2001).
- Lowe, S. W., Ruley, H. E., Jacks, T. & Housman, D. E. p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell* **74**, 957–967 (1993).
- Lowe, S. W. & Ruley, H. E. Stabilization of the p53 tumor suppressor is induced by adenovirus 5 E1A and accompanies apoptosis. *Genes Dev.* **7**, 535–545 (1993).
- Lowe, S. W., Jacks, T., Housman, D. E. & Ruley, H. E. Abrogation of oncogene-associated apoptosis allows transformation of p53-deficient cells. *Proc. Natl Acad. Sci. USA* **91**, 2026–2030 (1994).
- Lee, Y., Chong, M. J. & McKinnon, P. J. Ataxia telangiectasia mutated-dependent apoptosis after genotoxic stress in the developing nervous system is determined by cellular differentiation status. *J. Neurosci.* **21**, 6687–6693 (2001).
- Herzog, K. H., Chong, M. J., Kapsetaki, M., Morgan, J. I. & McKinnon, P. J. Requirement for Atm in ionizing radiation-induced cell death in the developing central nervous system. *Science* **280**, 1089–1091 (1998).
- Chong, M. J. *et al.* Atm and Bax cooperate in ionizing radiation-induced apoptosis in the central nervous system. *Proc. Natl Acad. Sci. USA* **97**, 889–894 (2000).
- Zaika, A., Irwin, M., Sansome, C. & Moll, U. M. Oncogenes induce and activate endogenous p73 protein. *J. Biol. Chem.* **276**, 11310–11316 (2001).
- Miyashita, T. *et al.* Tumor suppressor p53 is a regulator of *bcl-2* and *bax* gene expression *in vitro* and *in vivo*. *Oncogene* **9**, 1799–1805 (1994).
- Yin, C., Knudson, C. M., Korsmeyer, S. J. & Van Dyke, T. Bax suppresses tumorigenesis and stimulates apoptosis *in vivo*. *Nature* **385**, 637–640 (1997).
- Attardi, L. D. *et al.* PERP, an apoptosis-associated target of p53, is a novel member of the PMP-22/gas3 family. *Genes Dev.* **14**, 704–718 (2000).
- Oda, E. *et al.* Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science* **288**, 1053–1058 (2000).
- Jacks, T. *et al.* Tumor spectrum analysis in p53-mutant mice. *Curr. Biol.* **4**, 1–7 (1994).
- Yang, A. *et al.* p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature* **398**, 714–718 (1999).
- Yang, A. *et al.* p73-deficient mice have neurological, pheromonal and inflammatory defects but lack spontaneous tumours. *Nature* **404**, 99–103 (2000).
- Kinsella, T. M. & Nolan, G. P. Episomal vectors rapidly and stably produce high-titer recombinant retrovirus. *Hum. Gene Ther.* **7**, 1405–1413 (1996).
- Wells, J., Boyd, K. E., Fry, C. J., Bartley, S. M. & Farnham, P. J. Target gene specificity of E2F and pocket protein family members in living cells. *Mol. Cell. Biol.* **20**, 5797–5807 (2000).

Supplementary Information accompanies the paper on *Nature's* website (<http://www.nature.com>).

Acknowledgements

We thank J. Sage for critical reading of the manuscript, W. G. Kaelin and M. S. Irwin for human Tap73 α mammalian expression vector, L. Attardi for PERP cDNA, D. MacPherson for helpful discussions, and K. Olive for p53^{+/-} mice. This work was supported in part by the NIH and Howard Hughes Medical Institute. E.R.F. is a postdoctoral fellow of the Leukemia and Lymphoma Society of America. K.Y.T. is supported by the Medical Scientist Training Program and a Koch graduate fellowship. T.J. is an Associate Investigator of the Howard Hughes Medical Institute.

Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.J. (e-mail: tjacks@mit.edu).

Tissue culture and cell manipulation

Culture media, culture flasks and tissue processors.

InjectMan NI 2

Eppendorf www.eppendorf.com

Simple to use, from all angles

Eppendorf describes its new micromanipulator for microinjection applications as combining the advantages of a mechanical system with the benefits of a precisely functioning electromotive system. The device injects the microcapillary into the cell on an axial plane at any angle and withdraws it on the axial plane. It is said to be easy to operate and therefore suitable for less experienced users, while providing additional functions for those with more experience. A keypad with extra display guides operation, and an integrated clean/home function facilitates capillary handling and changing. Work levels and limits can be predetermined. The InjectMan NI 2 can be used with all commonly available inverse microscopes via a suitable adapter.



Eppendorf's InjectMan NI 2...

... makes microinjection simple.

ment of adherent mammalian cells and the surface is tested for monolayer formation using both cell lines and primary cells. Samples can be provided upon request to the manufacturer.

Biowave CO8000

WPA www.wpald.co.uk

A portable cell density meter

This hand-held device is a dedicated instrument for measuring the density of cells in suspension, useful in establishing when cells are ready to be harvested for protein expression, for instance. The CO8000 is portable, so the optical density of the culture can be measured next to the incubator for maximum convenience and accuracy. Up to 99 test results can be stored in the memory and then downloaded to a printer or PC. The Biowave cell density meter can be powered by mains electricity or by batteries. Little maintenance is required as the instrument uses an indefinite life LED and no filters. The CO8000 accepts a variety of cuvettes and tubes including 10, 12 and 16 mm diameter test tubes, and with no warm-up time. The results from the CO8000 are directly comparable to those obtained using a traditional spectrophotometer. An online demo is available at: <http://www.wpald.co.uk/co8000.html>

FBS brochure

Invitrogen www.invitrogen.com

A closer look at fetal bovine serum

Invitrogen's new Fetal Bovine Serum brochure describes the company's serum collection and processing methods, special serum variation-reducing programmes and services, and quality control tests. Complete

with country-specific disease data and including tables and graphs, the brochure is designed to assist scientists in identifying the serum most appropriate for their research. It contains a catalogue listing sera descriptions, applications and sources. The brochure also includes a 'frequently asked questions' insert with queries and reports on flocculence, heat inactivation, gamma irradiation and the thawing process. The brochure can be requested as a hard copy or online at the company's website.

HistoGene LCM kit

Arcturus www.arctur.com

A gentler approach

Optimized for the preparation of frozen tissue sections for laser capture microdissections (LCM), the HistoGene process features a proprietary stain solution designed to be less harsh than that of other protocols, thus preserving RNA integrity. The kit comes with reagents and supplies, including dehydration and staining solutions, slide jars and specially coated glass slides, and a detailed, step-by-step protocol.

Gentest

BD Biosciences www.bdbiosciences.com

Human hepatocytes, freshly delivered

These human hepatocytes from BD Biosciences are delivered by courier in flip-lock packaging to maximize viability and performance, with no film membranes, gel stabilizers or preservatives necessary. The hepatocytes are shipped in BD Hepato-STIM media on select lots of microplates and flasks. To register for this service, contact the company at hepatocytes_gentest@bd.com

EasyFlasks

Nunc A/S www.nuncbrand.com

Stacking up the benefits

EasyFlask cell culture flasks are now available in the full range of 25, 75 and 175 cm². The ergonomically designed screw cap opens with only a third of a turn and wide angled neck allows full access to the growth surface with pipette or cell scraper. The EasyFlask is designed to allow airflow under the flask to facilitate temperature equilibration on incubation even when stacked. The flasks have both printed and moulded volume graduations and are available with a patented filter cap or a vent/closed cap. Only virgin polystyrene complying with USP class VI is used in the production of the flasks. Nunclon treatment provides a surface suitable for attach-



Easy does it: cell culture flasks from Nunc.



Be prepared: Leica's microscopy sample processor.

EM TP

Leica www.leica-microsystems.com

This sleek model blows hot and cold

The Leica EM TP is a table-top tissue processor for use in resin processing of electron and light microscopy biological samples. It features a heater/cooler system, affording pre-heating and cooling from 4 to 60 °C to maintain a constant processing temperature. The instrument has a 99-program memory, variable agitation and a built-in ducted air ventilation exhaust system for safe handling of reagents. A membrane control panel allows additional programming to be input while a program is running, and includes a printer port, a battery back-up and a sealed, stainless steel chamber to protect the instrument from chemical spills. The reagent carousel holds 24 electron microscope or 12 light microscope 100 ml vials. A variety of baskets and capsules accommodate specimens of different types and sizes.

MACS 2002 catalogue

Miltenyi Biotec www.miltenyibiotec.com

An attractive assortment

This catalogue features MACS (magnetic cell sorting) products for the isolation of almost any cell type, magnetic cell isolation being completely compatible with flow cytometry. Applications include enumeration and isolation of viable antigen-specific T cells, and molecular biology procedures to isolate DNA, RNA and proteins, as well as their binding partners. Other offerings from Miltenyi Biotec include unique antibodies that enable labelling of previously unknown cell subsets such as those of blood dendritic cells and stem cells. The catalogue also highlights CE marked CliniMACS and CliniMACS^{plus} products for CD133⁺ (AC133⁺) and CD34⁺ stem and progenitor cell selection as well as isolation of CD14⁺ monocytes used to generate dendritic cells.

Immuno-trol Low Cells

Beckman Coulter www.beckmancoulter.com

Quality control at a lower level

This new control is an assayed, lysable, whole-blood quality control product for verification of reagent performance and methods used for staining and analysing targeted cells.

Designed for processing in the same manner as a whole blood sample, the control can be used throughout the immunophenotyping process, from sample preparation and antibody reactivity through to instrument processing and analysis. It offers reduced CD4 levels similar to specimens tested for immunodeficiency diseases. Fluorescence and light-scatter properties are similar to expected patterns. Other features include 90-day open-vial calibration stability, liquid ready-to-use reagents, cap-pierceable vials, and compatibility with most flow cytometry sample preparation systems. Used in conjunction with monoclonal antibody reagents and flow cytometry analysis, Immuno-trol is intended for diagnostic and clinical research laboratories conducting multicolour CD4 analysis.

B-339

Buchi Analytical www.buchi-analytical.com

Compact Kjeldahl distillation unit

Designed for high-speed steam distillation with integrated titration and evaluation, the B-339 Kjeldahl distillation unit comes with Buchi glassware and dedicated pumps for reagent dispensing. There are no dead volumes, preventing carry-over. The distillation process is visible and components can be reached through a hinged front panel. A software program is used to control the system, pre-select and store distillation and titration parameters, calculate results and print data. Applications include determination of nitrogen and protein by the Kjeldahl method and nitrogen determination in fertilizers using the Devarda method.

Proprotein convertases

Alexis Biochemicals www.alexis-corp.com

Technical bulletin

This update from Alexis describes the endoproteolytic processing enzymes called proprotein convertases. These have a specificity for basic motifs, such as RXX/R, RXXR and K/RR, in proteins in the secretory pathway, and are capable of generating bioactive proteins by limited endoproteolysis of the corresponding precursor proteins. Furin is the prototype of this family of processing enzymes. The bulletin also describes the company's monoclonal antibodies directed against major domains of human and mouse furin convertase.

High-resolution objective

Olympus Optical www.olympus-europa.com

New development in live cell observation

This high-resolution 20× objective has a numerical aperture of 0.95, allowing specimen observation to be carried out at different magnifications without changing the objective. Microscope adjustments, particularly changing objectives for observation at different magnifications, remains one of the major

causes of vibration. This objective overcomes that problem since magnifications can be changed from 7× to 80× without loss of resolution and with no need to move the objective. Its light collection performance facilitates observation deep within a specimen. The objective is particularly suitable for sensitive applications such as patch-clamping, gene injection into thick samples and membrane potential dye imaging. It is available with two new fixed-stage upright microscopes — the Olympus BX51WI and BX61WI.

These notes are compiled in the Nature office from information provided by the manufacturers.

Contacts

Publisher: Fabien Savenay
Editor: Paul Smaglik
Sales Director: Ben Crowe

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Manager:

Nevin Bayoumi (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)
Ben Corp (4974)
Andy Douglas (4975)
Laura Pearson (4977)
Netherlands/ Italy/ Iberia:
Evelina Rubio Hakansson (4973)
Scandinavia: Silje Opstrup (4994)

Production Manager:

Billie Franklin
To send materials use London
address above.
Tel +44 (0) 20 7843 4814
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

International

Advertising Coordinator:

Hind Berrada (4935)
Naturejobs web development:
Tom Hancock
Naturejobs online production:
Ben Lund

European Satellite Offices

France/ Belgium:

Christine Niox-Chateau
Tel + 33 (0) 1 43 20 16 51
Fax + 33 (0) 1 43 20 51 52
e-mail: c.nioxchateau@nature.com

Germany/ Austria/ Switzerland:

Patrick Phelan/ Kate Turner
Tel + 49 89 54 90 57 11/-2
Fax + 49 89 54 90 57 20
e-mails: p.phelan@nature.com
k.turner@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Director:

Ben Crowe
US Sales Manager: Peyton Mason

US Advertising Coordinator:

Ashly de Leon

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatamachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746
e-mail: k.cowan@naturejp.com
Japan Manager: Kate Cowan

naturejobs

French disconnection

At an informal meeting with students at the Paul Sabatier University in Toulouse last month, I asked who among them wanted to gain a PhD in science. The majority put up their hands. Then I asked which of them wanted to work in academia, and most of the hands stayed in the air. Finally, I asked who wanted to stay in France — the show of hands remained largely unchanged.

Unfortunately for those students, sticking to these goals could greatly limit their long-term career prospects. In France, as in much of Europe, there is a dearth of permanent positions, and the young scientists competing for these posts are becoming increasingly frustrated (see *Nature* **414**, 145; 2001).

One of the main problems in France is that the government offers little in the way of graduate or postdoctoral funding. Many of the scientists I met in Toulouse told me that they relied on the Human Frontier Science Program and Europe's Marie Curie fellowships for their postdocs. But neither of these programmes offers much solace to French students seeking academic posts within their own country — both stipulate that the postdocs they fund must come from abroad.

So how can French students improve their prospects? The first step might be to broaden their search to include industrial postdocs. They could also consider widening their options by seeking postdoc positions in other European countries. Then, having increased their skills and experience, they might be able to return home thanks to one of a variety of available repatriation schemes (see *Naturejobs* 5; 28 March 2002).

It might seem bizarre to French students that the way to secure work at home is to leave the country. But such a path may well be the best bet for them — and for many other Europeans.

Paul Smaglik

Naturejobs editor



Contents

CAREERS AND RECRUITMENT

Researchers marshal their forces in the fight against cancer

p4

WWW.NATUREJOBS.COM

Career centre
Information on the scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

SCIENTIFIC ANNOUNCEMENTS

SCIENTIFIC EVENTS



Martin Abelloff (below) sees the building that houses the Sidney Kimmel Comprehensive Cancer Center (bottom) as a 'medical school in microcosm'.



Tearing down the barriers

In the fight against tumours, comprehensive cancer institutes are deploying a wide range of different strategies in an attempt to encourage the sharing of ideas and materials among researchers, says Steve Bunk.

Among cancer researchers, the best way to confront the complexities of disease is by cooperating. But the exchange of information between different specialists does not always happen spontaneously. It can require inventiveness in the use of physical space, in structuring teams, even in decisions about who should set research objectives and monitor progress.

Comprehensive cancer centres — which combine basic, population and clinical research — have long experience in such strategies. But, just like the humans who they aim to marshal, these centres both in the United States and Europe show constant flux and wide differences in the way that they are organized.

"One size doesn't fit all," says Martin Abelloff, director of the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins in Baltimore, Maryland. His centre uses a range of methods to bring people together. For example, a few pathologists have integrated their laboratories into oncology — outside of the formal pathology department.

Others belong to working groups that might focus on either a specific cancer or approach, such as signal-transduction inhibition or bone-marrow transplantation. These groups often include many of the skills needed to study, diagnose and treat cancers, among them molecular biology, experimental therapeutics, medical oncology, radiation, surgery and pathology.

Some team members are from the medical school,

nursing school or the public-health campus, all nearby. Central to the development of interdisciplinary understanding among them is frequent face-to-face contact, Abelloff says. "Space, in any academic institution, is probably the most precious resource," he explains (see 'Open plans' opposite).

Two years ago, the cancer centre's 440 employees moved into a new building that Abelloff likens to a medical school in microcosm. It consolidates the labs and offices of basic, biostatistical and clinical researchers, shared resources and conference rooms.

Abelloff thinks that one of the biggest challenges is establishing a thorough recognition among lab-based investigators of the numerous subtle problems that can arise when a promising molecular therapy that has been tested on animals is applied to humans. This includes the nuances of clinical-trial design, which can affect decisions made during preclinical research.

Accordingly, clinicians must also appreciate details of bench science. "When it really works, I've been in sessions where I can't identify a person as a PhD or an MD, as lab-based or clinically based, because they really are speaking the same language," Abelloff says.

COMBINED EFFORT

In a similar way, the Fred Hutchinson Cancer Research Center in Seattle seeks cross-fertilization through a combination of egalitarianism and personal freedom. "The basic science division has no programmes per se," says division director Mark Groudine, a radio-oncologist at the University of Washington, Seattle. "Everyone works on a model they want to work on. We just have really good people and turn them loose."

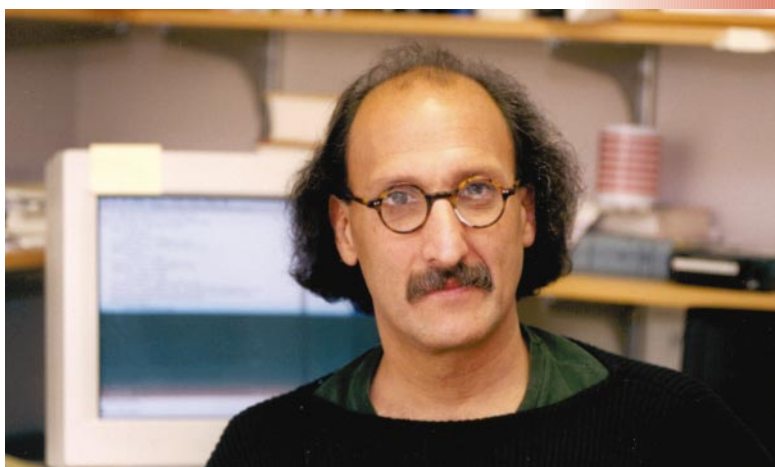
The Hutchinson centre discourages academic empire-building by allocating space based solely on rank rather than on achievements or grant monies. Even salaries are equalized by rank and time-plus-rank, using models generated by the centre's biostatisticians.

Discretionary funds within the centre tend to go towards junior people's research or to any lab in financial straits, Groudine says. The result of these and other non-hierarchical policies is "an enormous sharing of ideas and reagents in this division".

Centre-wide, the top leadership meets every other week, in part to cultivate a viewpoint that favours the good of the centre over the isolation of resources in specific departments. For example, the clinical division recently provided some of its lab space to help with a major recruitment drive for solid-tumour programmes, which division directors had identified as a top priority.

Groudine says that the basic, clinical and public-health sciences could each be a free-standing institute. Every scientist in the newest division, human biology, also has an appointment in another division, to help increase cross-fertilization of ideas.

To exchange and gather information from staff scientists, the centre provides an array of clubs, retreats, symposia, mini-courses, lectures and seminars — even a programme for postdocs and graduate students to be mentored by someone from a different discipline. Groudine says that the sort of personalities who fit best into the centre are: "Bright, interactive people who like being in an atmosphere where there's enormous collaboration."



Mark Groudine espouses the collaborative spirit.

What a given researcher might or might not want appears to be important to success in a comprehensive cancer center. Waun Hong, who heads the 13-department Division of Cancer Medicine at the University of Texas M. D. Anderson Cancer Center in Houston, knows what personal traits he would seek in choosing between two job candidates with the same technical skills. "I'd select the person who has more character, is more friendly, is collaborative; the unselfish person who is willing to share and give," he says.

Waun Hong (left) is seeking strong characters.

NOVEL APPROACH

Hong says that research in a comprehensive cancer centre is quite unlike the traditional model of independent study driven by a principal investigator. Earlier in his career, a clinical trial might have consisted of checks for toxicity and response in 30 or 40 patients, he says. Nowadays, such studies extend to mechanisms of action, pharmacokinetics, cellular markers, imaged measurement of responses to identify predictive factors and biostatistical correlation of these multiple markers with clinical responses.

Also needed are a tissue bank, serum bank, patients, and surgeons to do biopsies. Other capabilities that come from outside Hong's division include molecular imaging, pathology and biostatistics. "The process is incredibly complicated. That's why it's difficult to put numbers on the scoreboard," he says. A narrowly



Open plans

At the European Institute of Oncology in Milan, the concept of a laboratory without walls has been taken to an extreme. The institute, which includes clinical facilities and 226 patient beds, divides its research activities into a Division of Epidemiology and Biostatistics and a Department of Experimental Oncology.

Seven years ago, the space for experimental oncology was redesigned to replace assigned

rooms with a specific number of benches or workstations allocated to each group leader. These stations are scattered throughout the department, eliminating the proximity of team members in favour of putting people from different specialities side-by-side.

"The minor drawback of having people from the same group in different areas is largely overcome by the great number of exchanges and the

generation of new ideas among people from different groups," says the institute's research coordinator, Giuseppe Della Porta. He acknowledges that group leaders had to be coaxed into giving up the idea of 'their space', and adds that new ways of calculating and handling budgets were also needed.

Centralization of facilities is another way that departmental 'coalescence' is promoted, he adds. This extends beyond the

usual centralized facilities, such as DNA sequencing, imaging and microarrays. All solutions and media are prepared in one place, and the performance of microinjections, molecular pathology and bioinformatics is similarly centralized. The result is lower costs, higher quality and improved technological development, fostering a strong common interest within the department.

S.B.



Mix and match: at the M. D. Anderson Cancer Center, a multidisciplinary approach to research is encouraged.

defined investigation won't suffice. "You need to have a critical mass and a team effort."

Nevertheless, some of the centre's scientists spend all of their time doing individual research and some are entirely focused on clinical work, says Leonard Zwelling, vice-president for research administration. Multidisciplinary groups sometimes rely on departments to do lab work but Zwelling prefers to see basic scientists on those teams. He says that it improves access to tissues and to patient information, which increases the clarity of researchers working with particular disease entities.

For example, everyone on the team learns about tumour behaviour, metabolic potentials and angiogenic effects. When the centre's new building is finished in about two years, Zwelling says that each pathology department will be physically grouped with its closest-linked research programme: for example, breast pathology with breast cancer.

Many of the centre's scientists do not hold teaching positions with the affiliated University of Texas Health Science Center at Houston across the street, but the focused mission of cancer research brings a constant influx of fellows and postdoctoral trainees in the biomedical sciences. "By definition, we're almost advanced training," Zwelling says. There's no part of the centre's skill set that isn't in demand, he affirms.

EUROPEAN PERSPECTIVE

The Institute of Cancer Research (ICR) in London, which together with the neighbouring Royal Marsden Hospital makes up Europe's largest comprehensive cancer centre, doesn't expect all of its people to have a range of personal skills in such spheres as communication, lateral thinking or teamwork, says chief executive Peter Rigby. "Some team leaders are highly successful because they are completely focused on their own area," he says. "In the end, we pick the best people in the particular area. It generally works out that enough of them have the other skills that the institution works together as a whole."

Structurally, he sees the ICR as a hybrid, an academic institution with a clearly defined mission

that drives its hiring policy. He admits that this sets up a tension between academic and mission-oriented research, which the centre is now trying to address. The clearest case of this tension, he adds, is probably the CRC Centre for Cancer Therapeutics in Sutton, which is organized like a biotech company, although team leaders are evaluated using academic criteria.

Keith Willison, head of the ICR's Chester Beatty Laboratories in London, says that the institute followed a classic academic model until recently. But now, instead of simply hiring the smartest people available and providing them with resources, the ICR has a faculty structure. Faculty members are independent, and are the only ones allowed to apply for grant funding or supervise PhD students.

GOING FOR GOALS

But a major strand of the research strategy pursued by the ICR and the Royal Marsden Hospital is a highly goal-orientated drug-discovery programme. The lead investigators are faculty members. "Senior scientists are trying to find techniques for progressing our scientific strategy without undermining scientific independence," he says.

One such effort, Willison adds, was the recent establishment of a section to elucidate high-resolution structures of proteins involved in cell signalling and cancer. Setting up this section cost the institute about £2.5 million (US\$3.5 million), but there are now 40 scientists who have solved important structures and published their results in top journals. "This is highly academic work but it also vital for our structure-based drug-discovery programme," Willison says.

Daniel Louvard faced a similar problem seven years ago, when he became research-division director at the Curie Institute in Paris. "I realized a lot of skills were not being used optimally, because people were not working together," he says. He responded by creating an incentive programme for cooperation, which makes



The drive for cancer therapies is shifting the emphasis of research onto cooperation.

internal funds available for consumables in research projects that bridge chemistry and biology, physics and biology, or basic and clinical biomedical sciences.

Projects co-led by scientists from each of the two disciplines follow a well-defined theme and are funded for four years with a possible one-year extension, a longevity that Louvard says is "very unusual in France". Among the investigation themes are antitumour immunotherapy, micrometastasis and genotoxicology. Louvard believes that the programme is helping to create new ways of understanding complex phenomena, which he says will be increasingly important to researching the complexities of cancer. "What's exciting is not what's happening in the middle of the football field, it's what's happening at the edges," he says. "That's where the creativity, vision and surprises are."

Central to such discoveries is bioinformatics, which Louvard says is easily the institute's most sought-after

skill, an assessment seconded by several other cancer-centre leaders. The French government is trying to help to address this shortage by initiating a cooperative genomics venture this year between several organizations, including the Curie Institute and France's only other comprehensive cancer centre, the nearby Gustave-Roussy Institute. They will combine some of their funds to bridge technological gaps in chromosomal mapping and microarrays, and will form a "virtual network", he says.

BUILDING BRIDGES

The Curie Institute's research division has about 600 employees, whereas the medical division includes a 200-bed hospital and 800 personnel. Organized as a foundation, the institute provides equipment, facilities and financial support for internal grants and various special programmes for about 150 graduate students and 100 postdocs, but not the salaries for its 200 permanent scientists in the research division. They are hired by either of two federal bodies, the CNRS — France's main research agency, or the National Institute for Health and Medical Research (INSERM).

All 12 labs, which are like small departments, are evaluated every four years by government and institute committees. The labs could be disbanded for poor performance or moved out of the centre if researchers stray too far from the overall mission. But in the latter case, Louvard says: "The intellectual environment, the quality of the facility, and the internal support are so attractive that nobody does that."

Comprehensive cancer centres attest to the many different ways in which high-quality research can be achieved. These differences arise from one common trait — flexibility in organizational structure. As Johns Hopkins' Abeloff points out, flexibility is the way to accommodate individuality. One of his biggest challenges is dealing with specialists who have long, successful track records but who still show limited understanding of other research areas. This often stems from traditional, principal-investigator-driven research, he says, but he is quick to defend this model if it produces good, fundamental research.

"It's a given that most researchers want to have an impact on cancer," Abeloff says. "What isn't a given is that collaborations just naturally occur." He acknowledges that bringing everyone together probably isn't even a good idea. Cooperation, it appears, is a sum greater than its parts.

Steve Bunk is a science writer based in Alexandria, Virginia.

Web links

Kimmel Cancer Center
 ♦ www.hopkinscancercenter.org
 Fred Hutchinson Cancer Research Center
 ♦ www.fhcrc.org
 M.D. Anderson Cancer Center
 ♦ www.mdanderson.org
 Institute of Cancer Research
 ♦ www.icr.ac.uk
 Curie Institute
 ♦ www.curie.fr
 Cancer Research UK
 ♦ www.cancerresearchuk.org

Money matters

The launch this February of Cancer Research UK as the world's largest cancer charity could do much for jobs, but it cannot compensate for what joint general director Gordon McVie calls the UK government's "massive underfunding" of cancer studies.

Arising from a merger of the Cancer Research Campaign and the Imperial Cancer Research Fund, the new charity can summon some £150 million (US\$214 million) a year in funding for cancer research among its team of 3,000 scientists, doctors and nurses in British institutions, says McVie.

In contrast, the government's Medical Research Council provides less than £60 million annually for research, McVie says. Britain's other cancer charities add another £40 million, but industry is the big contributor, providing about £500 million per year.

With its arrival, Cancer Research UK announced a £75-million investment initiative

that will include improvements to several comprehensive cancer centres. Britain has such centres in London, Leeds, Oxford, Cambridge, Birmingham, Glasgow and Edinburgh.

Cancer Research UK has signed a new contract with the US National Cancer Institute (NCI) to exchange resources, information and personnel, but the NCI's \$3.5-billion budget is tough competition.

The NCI has designated 41 cancer centres nationwide as 'comprehensive', which means that each displays depth and breadth of research in basic, clinical and population sciences.

They also offer interactive research that aims to act as a bridge between these areas. The institute's 2003 budget request for research and infrastructure expenditures under its cancer centres programme is \$81.3 million. S.B.



Gordon McVie: anticipates a bright future for Cancer Research UK.